

Inkorrektheitsphänomene und Regularisierung bei der Parameterschätzung für Jump-Diffusions-Prozesse

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium

(Dr. rer. nat.)

TECHNISCHE UNIVERSITÄT CHEMNITZ

Fakultät für Mathematik



TECHNISCHE UNIVERSITÄT
CHEMNITZ

vorgelegt von Dipl.-Math. oec. Dana Düvelmeyer
geb. am 20. 08. 1977 in Greiz

Chemnitz, den 20. Januar 2005

Inhaltsverzeichnis

1	Einleitung	4
2	Jump-Diffusions-Prozesse	8
2.1	Kursmodell	11
2.2	Eigenschaften der Returns	17
2.2.1	Stationarität der Returns	17
2.2.2	Verteilung der Returns	19
2.3	Spezialfall geometrische Brownsche Bewegung	25
3	Parameterschätzung – das inverse Problem	27
3.1	Übergang von Kursdaten zu Returndaten	27
3.2	Statistischer Zugang zur Parameterschätzung	33
3.2.1	Maximum-Likelihood-Methode	34
3.2.2	Momentenmethode	37
3.3	Analytischer Zugang	38
3.3.1	Modifikation des Abstandsbegriffs	41
4	Eigenschaften des Vorwärtsoperators	50
4.1	Stetigkeit des Operators	51
4.2	Injektivität des Operators und Stabilitätsbetrachtungen	61
4.3	Grenzsituationen	70
5	Inkorrektheitsphänomene	74
5.1	Diskretisiertes Problem	75
5.2	Verwendete Optimierungscodes	77
5.3	Numerische Fallstudie	77
5.3.1	Verwendete Daten	80

5.3.2	Ergebnisse im ungestörten Fall	81
5.3.3	Ergebnisse im leicht gestörten Fall	83
6	Ein Multiparameter-Regularisierungszugang	86
6.1	Zusatzinformation durch Semiinvarianten	86
6.2	Regularisierungsalgorithmus	88
6.2.1	Eine Fixpunktiteration	92
6.2.2	Numerisches Beispiel	94
6.3	Numerische Modifikation	97
6.3.1	Fallstudie zum Einfluss des Exponentialparameters	98
6.4	Vergleichsrechnung zur Maximum-Likelihood-Methode	102
7	Anwendungen in der Overreaction-Theorie	108
7.1	Ein Overreaction-Indikator	109
7.2	Maße für die Gegenbewegung	112
7.3	Erwartungswerte der Laufzeitextrema	114
7.3.1	Spezialfall der geometrischen Brownschen Bewegung	115
	Abkürzungs- und Symbolverzeichnis	121
	Thesen	127

Kapitel 1

Einleitung

Häufig werden mathematische Modelle genutzt, um komplexe reale Prozesse aus Natur, Technik oder Wirtschaft nachzubilden. Die Modelle zur Beschreibung der Kursentwicklung von Finanztiteln (Assets) basieren üblicherweise auf stochastischen Prozessen, die durch Modellparameter spezifiziert werden. In vielen finanzmathematischen Anwendungen müssen diese Parameter an empirische Assetpreise angepasst werden, damit das verwendete Modell eine gute Approximation des viel komplexeren realen Prozesses darstellt. Der Anpassung (Kalibrierung) des Modells an gegebene Marktdaten kommt eine zentrale Bedeutung zu. Falls die geschätzten Parameter keine gute Approximation der unbekannten exakten Parameter darstellen, können von dem Modell abgeleitete Größen fehlerhaft sein. Werden beispielsweise Preise von Derivaten in einem Modell hergeleitet, so ist es bei ungünstiger Parameterwahl durchaus möglich, dass trotz richtiger Formeln die berechneten Preise der Derivate stark von den in der Realität beobachteten Größen abweichen.

Die vorliegende Dissertation widmet sich dem Problem der Bestimmung der Parameter, die in dem Kursmodell von Jump-Diffusions-Prozessen auftreten. Dabei steht ein Fünf-Parameter-Modell im Mittelpunkt der Betrachtungen. In der Literatur werden verschiedene statistische Zugänge zur Parameterbestimmung diskutiert. Es gibt Arbeiten, die auf Stabilitätsprobleme bei gewissen Parameterkonstellationen hinweisen, sofern die Parameter mit der herkömmlichen Maximum-Likelihood-Methode geschätzt werden. Numerische Rechnungen zu statistischen Standardverfahren haben gezeigt, dass diese Stabilitätsprobleme bei viel mehr als den dort diskutierten Parameterkonstellationen und insbesondere dann auftreten, wenn die Parameter aus einer relativ kleinen Zahl beobachteter Assetpreise bestimmt werden. Daher untersuchen wir das Problem der Parameterschätzung in dieser Arbeit unter Einbeziehung von Methoden aus der Theorie inverser Probleme, da deren zentrales Anliegen die Analyse und Regularisierung inkorrektur und instabiler inverser Aufgaben ist. In dieser Arbeit werden Phänomene der Instabilität der Parameterbestimmung herausgearbeitet und analysiert. Hierfür leiten wir eine entsprechende nichtlineare Operatorgleichung $A(p) = f$ her, indem wir den Zusammenhang zwischen einer von den Parametern \underline{p} abhängigen Trajektorie des Jump-Diffusions-Prozesses und der Dichtefunktion f der Returns nutzen. Diese Operatorgleichung untersuchen wir bezüglich ihrer Korrektheit. Wir zeigen anhand einer

Fallstudie mit simulierten Daten, dass bei der numerischen Lösung Inkorrekttheitsphänomene auftreten, sobald die Daten mit kleinen Datenfehlern behaftet sind. Da wir an einem stabilen Schätzverfahren interessiert sind, ist ein weiteres Ziel dieser Arbeit die Regularisierung des inversen Problems der Parameteridentifikation.

Dementsprechend ist die Arbeit gegliedert: Zunächst betrachten wir im 2. Kapitel das Modell der Jump-Diffusions-Prozesse. Wir führen die stochastische Differentialgleichung ein, die das Kursverhalten beschreibt. Das hier betrachtete Kursmodell wird durch insgesamt fünf Parameter $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)$ spezifiziert. Die Parameter μ und σ bestimmen Drift und Volatilität des Diffusionsteils. Die Sprungintensität λ beeinflusst die Häufigkeit der Kurssprünge während die Parameter μ_Y und σ_Y die mittlere Höhe der Sprünge sowie deren Volatilität bestimmen. Von der Lösung der stochastischen Differentialgleichung leiten wir die Returns ab, zeigen deren Stationarität und berechnen deren Verteilungs- und Dichtefunktion sowie die Momente und Semiinvarianten. Sofern die Sprungkomponente das Kursverhalten nicht beeinflusst, ergibt sich als Spezialfall der Jump-Diffusions-Prozesse das Modell der geometrischen Brownschen Bewegung. Die Sprungkomponente verschwindet einerseits, wenn die Sprungintensität λ null beträgt und andererseits, wenn sowohl die mittlere Sprunghöhe μ_Y als auch die Sprungvolatilität σ_Y null sind.

In Kapitel 3 führen wir anschließend das inverse Problem der Parameterschätzung ein. Die empirischen Returns, die aus einer Zeitreihe von Aktienkursen berechnet werden, bilden die Grundlage für alle Schätzverfahren. Zunächst betrachten wir die statistischen Standardzugänge zur Parameteridentifikation für das Kursmodell der Jump-Diffusions-Prozesse und diskutieren eine in der Literatur vorgestellte Modifikation der Maximum-Likelihood-Methode. Anschließend stellen wir einen analytischen Zugang zur Parameterschätzung vor und definieren den Vorwärtsoperator A des direkten Problems, der den Parametervektor \underline{p} des Jump-Diffusions-Prozesses die Dichtefunktion f der Returns zuordnet.

Die Eigenschaften des Vorwärtsoperators untersuchen wir im 4. Kapitel. Wir zeigen die Stetigkeit des Operators, der vom Raum \mathbb{R}^5 in den Raum $L^2(\mathbb{R})$ abbildet. Außerdem diskutieren wir die Injektivität des Operators und die Stabilität der inversen Abbildung. Wir zeigen, dass der Vorwärtsoperator nur für Parametervektoren, für die sich der Spezialfall der geometrischen Brownschen Bewegung ergibt, nicht injektiv ist. Daher ist jede Lösung $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)$ der Operatorgleichung $A(\underline{p}) = f$ mit $\lambda(\mu_Y^2 + \sigma_Y^2) \neq 0$ eindeutig bestimmt. Ferner zeigen wir, dass für alle Lösungen mit $\lambda(\mu_Y^2 + \sigma_Y^2) = 0$ die Diffusionsparameter μ und σ übereinstimmen. Da alle Parameter $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)$ mit $\lambda(\mu_Y^2 + \sigma_Y^2) = 0$ auf den Spezialfall der geometrischen Brownschen Bewegung führen, sind die Werte der Sprungparameter λ , μ_Y und σ_Y bedeutungslos. Wir analysieren ausführlich die Korrektheit der Operatorgleichung $A(\underline{p}) = f$. Durch Beschränkung der Sprungparameter $\lambda \leq \lambda_{\max}$, $|\mu_Y| \leq \mu_{Y\max}$ sowie $\sigma_Y \leq \sigma_{Y\max}$ können wir zeigen, dass die Konvergenz einer Folge f_n im Raum $L^2(\mathbb{R})$ bzw. $C(\mathbb{R})$ gegen die Dichtefunktion f_0 die Konvergenz der Diffusionsparameter der Urbilder $\underline{p}_n \in U(f_n)$ bezüglich des Operators A gegen die Diffusionsparameter aller Parametervektoren $\underline{p}_0 \in U(f_0)$ nach sich zieht. Ferner zeigen wir, dass der Ausdruck $\lambda_n(\mu_{Y_n}^2 + \sigma_{Y_n}^2)$ mit den Sprungparametern λ_n , μ_{Y_n} und σ_{Y_n} der Urbilder $\underline{p}_n \in U(f_n)$ für $n \rightarrow \infty$ gegen den Ausdruck $\lambda_0(\mu_{Y_0}^2 + \sigma_{Y_0}^2)$

konvergiert. Folglich können kleine Datenfehler keine Stabilitätsprobleme verursachen, sofern die Sprungparameter beschränkt werden. Wir betrachten für jeden Parameter die Grenzsituationen. Die Fälle $\sigma \rightarrow 0$ und $\lambda \rightarrow \infty$ erweisen sich dabei als besonders interessant und werden ausführlich analysiert. Falls die Sprungintensität λ nicht beschränkt wird, tritt eine asymptotische Nichtinjektivität auf, die Instabilitätsprobleme verursachen kann.

Im 5. Kapitel illustrieren wir anhand einer umfangreichen numerischen Fallstudie, dass trotz der theoretischen Stabilitätsaussagen bei der numerischen Lösung Inkorrektheitsphänomene auftreten. Wir konstruieren leicht gestörte Datenfunktionen und zeigen, dass bei der numerischen Lösung des Least-Squares Problems, welches sich aus der diskretisierten Operatorgleichung ableitet, mitunter erhebliche Abweichungen zwischen den geschätzten und exakten Parametern auftreten.

Zur Überwindung der Inkorrektheitsphänomene stellen wir anschließend in Kapitel 6 einen Multiparameter-Regularisierungszugang vor. Es erweist sich, dass die Semiinvarianten genügend Zusatzinformationen zur Regularisierung enthalten. Bei der von uns vorgeschlagenen Regularisierung wird das Minimum des ursprünglichen Least-Squares Problems nicht mehr für alle Parametervektoren \underline{p} aus dem Definitionsbereich D bestimmt, sondern nur für Parametervektoren deren Semiinvarianten sich in der Umgebung der empirischen Semiinvarianten befinden. Der Regularisierungszugang führt auf ein nichtlineares und nichtkonvexes Optimierungsproblem, dessen Optimallösungen geeignete Approximationen des exakten Parametervektors sind. Wir führen ein zu diesem Problem verwandtes Problem ein, dessen Regularisierungsfunktional in Analogie zur Tichonov-Regularisierung gewählt wurde. Hier steuert man die Stabilisierung allerdings nicht mit einem skalaren Regularisierungsparameter, sondern mit einem Regularisierungsparametervektor. Wir leiten eine Fixpunktiteration her, mit der die Optimallösung näherungsweise berechnet werden kann, indem wir für verschiedene Regularisierungsparametervektoren jeweils die Lösung des verwandten regularisierten Problems berechnen. Wir beweisen, dass jeder Fixpunkt, gegen den dieser Algorithmus konvergiert, einer Optimallösung entspricht. Wir illustrieren anhand zahlreicher numerischer Rechnungen mit simulierten Daten, dass der Algorithmus in praktischen Situationen konvergiert und die erhaltenen Lösungen sehr gute Approximationen der exakten Lösung darstellen. Wir stellen verschiedene Modifikationen zu diesem Algorithmus vor, um die benötigten Iterationsschritte zu verringern und diskutieren das Problem der Wahl numerischer Parameter anhand einer Fallstudie. Am Ende dieses Kapitels simulieren wir Kursdaten und ermitteln die Parameter einerseits mit der Maximum-Likelihood-Methode und andererseits mit dem vorgestellten Multiparameter-Regularisierungszugang. Mit diesem Vergleich illustrieren wir, dass wir mit unserer Methode eine Stabilisierung des Problems der Parameteridentifikation erreichen können.

Zum Abschluss präsentieren wir im 7. Kapitel eine konkrete Anwendung der Parameterschätzung in Jump-Diffusions-Prozessen. Wir betrachten einen Indikator zur Erkennung von Overreaction-Effekten, der auf der Berechnung von Wahrscheinlichkeiten extremer Kursbewegungen beruht. Der Overreaction-Effekt wird durch irrationales Handeln vieler Marktteilnehmer in ungewöhnlichen und unerwarteten Situationen verursacht. Entsprechend sollte das Auftreten solcher Überreaktionen und die damit ver-

bundenen extremen Kursbewegungen sehr unwahrscheinlich sein. Daher identifizieren wir eine Overreaction-Situation, wenn die Kurse an drei aufeinander folgenden Tagen gestiegen oder gefallen sind und die Wahrscheinlichkeit, dass solch eine Kursbewegung eintritt sehr klein ist. Grundlage für eine zuverlässige Funktionsweise des Overreaction-Indikators ist die genaue Berechnung der Wahrscheinlichkeiten, die wiederum von der Genauigkeit der geschätzten Parameter abhängt. In der Literatur zum Overreaction-Effekt wird die These überprüft, dass extremen Kursbewegungen entsprechende Gegenbewegungen folgen. Wir entwickeln Maße, die die Gegenbewegung einer Overreaction-Situation nach einer entsprechenden Indikation des Overreaction-Indikators bewerten. Weiterhin vergleichen wir diese Maße mit den Erwartungswerten der Laufzeitextrema der Returns und erhalten somit Aufschluss über das Ausmaß der Gegenbewegung. Schließlich stellen wir einen allgemeinen Zugang zur Berechnung dieser Erwartungswerte für das Kursmodell der Jump-Diffusions-Prozesse vor und leiten für den Spezialfall der geometrischen Brownschen Bewegung explizite Formeln her.

Abschließend möchte ich mich bei allen herzlich bedanken, die zum Gelingen dieser Dissertation beigetragen haben. Besonderer Dank gebührt meinem Doktorvater, Herrn Prof. Dr. Bernd Hofmann, für seine hervorragende, unermüdliche Betreuung und die zahlreichen Diskussionen, Anmerkungen und Anregungen. Seine Begeisterung für die Theorie inverser Probleme hat auch meine Faszination für dieses interessante Aufgabenfeld geweckt. Trotz hoher Belastung hat er immer Zeit gefunden, die neueste Version der Arbeit zu lesen und ausgesprochen konstruktiv zu kritisieren. Ebenso herzlich möchte ich mich bei Herrn Prof. Dr. Ralf Wunderlich bedanken, der trotz vollen Terminkalenders mit zahlreichen konstruktiven Anmerkungen und Verbesserungsvorschlägen wesentlich zum Gelingen dieser Arbeit beigetragen hat. Auch den Mitgliedern der Arbeitsgruppe Stochastik mit Herrn Prof. Dr. Jürgen vom Scheidt an der Spitze gebührt mein Dank für die große Diskussionsbereitschaft und Unterstützung bei stochastischen Fragestellungen. Nicht zuletzt möchte ich mich bei meinem Freund Marcel Uhlig, meiner Familie und meinen Freunden bedanken, die mich immer tatkräftig unterstützt haben. Danke, dass Ihr so manche geistige Abwesenheit, schlechte Laune und Zeitdefizite ertragen habt und mir immer einen Ausgleich zur wissenschaftlichen Arbeit gegeben habt.

Kapitel 2

Jump-Diffusions-Prozesse

In der Finanzmathematik nutzt man mathematische Modelle, um die Kursentwicklung eines Assets zu modellieren (siehe z.B. [2], [43], [59], [61], [65], [67]). Da das Kursverhalten eines Wertpapiers im Allgemeinen nicht sicher vorhersagbar ist, besitzen diese Modelle eine stochastische Komponente. Der Kurs des Wertpapiers wird daher durch einen stochastischen Prozess $(S_t, t \in I)$ über dem Zeitintervall $I \subseteq \mathbb{R}_+$ sinnvoll modelliert. Auch wenn das Kursverhalten nicht prophezeibar ist, weisen die Kurse doch systematische Verhaltensmuster auf. Um die Vielfalt der verschiedenartigen Assets möglichst realistisch zu beschreiben, kann man unterschiedliche Modelle nutzen, die das jeweilige Muster charakterisieren. So werden beispielsweise Mean-Reversion-Prozesse (z.B. [71]) zur Modellierung von Zinsentwicklungen genutzt, da Zinssätze sich mittelfristig immer wieder ihrem Mittelwert annähern. Das Standardmodell zur Beschreibung von Aktienkursen ist das der geometrischen Brownschen Bewegung, welches im Black-Scholes Modell [8] auch zur Beschreibung des Underlyings bei der Bestimmung von Optionspreisen genutzt wird. Eine Annahme dieses Kursmodells ist, dass die zukünftigen Aktienkurse nur von dem derzeitigen Kurs nicht aber von der Wertentwicklung der Vergangenheit abhängen. Das Modell der geometrischen Brownschen Bewegung impliziert, dass die Pfade des stochastischen Prozesses $(S_t, t \in [0, \infty))$ stetig sind, d.h. die Aktienkurse sich stetig verändern, und die logarithmierten Returns $r_{s,t} = \ln \frac{S_t}{S_s}$ normalverteilt sind.

Betrachtet man am Kapitalmarkt die tatsächlich beobachteten Returns, so haben diese mitunter leptokurtische Verteilungen. Die Returns besitzen in praxi im Vergleich zur Normalverteilung um den Mittelwert herum und an den Rändern mehr Wahrscheinlichkeitsmasse. In dem Buch [12, S. 209] findet man eine Übersicht über die statistischen Eigenschaften von Aktienkursen und Aktienindizes, die in verschiedenen empirischen Studien der letzten Jahre festgestellt wurden. Viele solcher Studien zeigen recht deutlich, dass das Modell der geometrischen Brownschen Bewegung, welches normalverteilte Returns impliziert, empirisch beobachtete Eigenschaften nicht ausreichend gut beschreiben kann. In der Literatur gibt es daher viele Modifikationen des Kursmodells der geometrischen Brownschen Bewegung, um diese empirisch beobachteten Eigenschaften besser beschreiben zu können.

Treten am Markt gewöhnliche Informationen auf, die „normale“ Kursänderungen hervorrufen, sind stetige Kursveränderungen mit der Realität gut vereinbar. Allerdings

können auch ungewöhnliche Informationen auftreten, die dann teilweise sprungartige Kursveränderungen hervorrufen¹. Derartige Kurssprünge findet man in der Praxis recht häufig bei der Betrachtung von Aktienkursen und von Intradaykursen bei Aktienindizes (siehe Abbildung 2.1). Aber man kann sie auch in modifizierter Form bei Schlusskursen von Aktienindizes beobachten (siehe Abbildung 2.2).

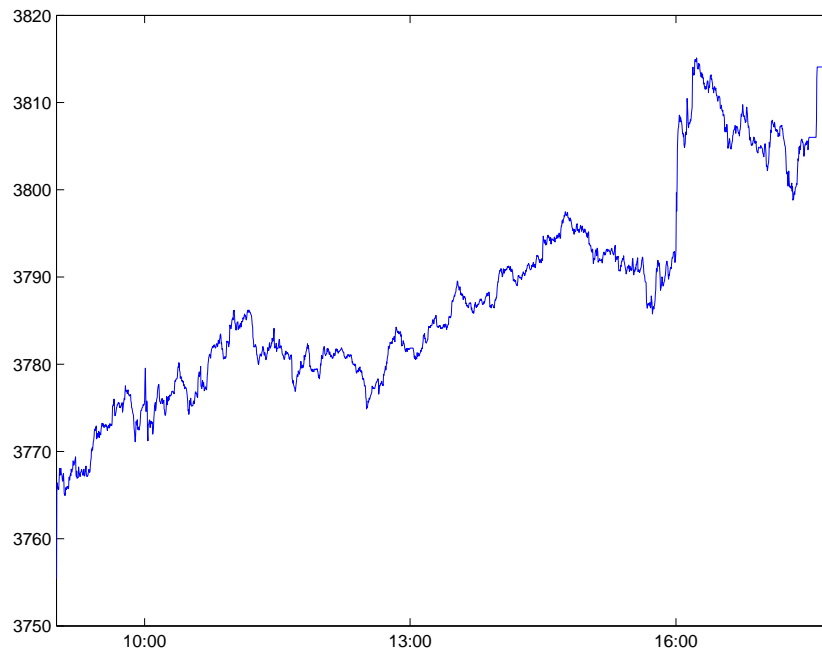


Abbildung 2.1: DAX am 27.07.04

Wir betrachten daher in dieser Arbeit das Kursmodell der Jump-Diffusions-Prozesse, mit dem einerseits Kurssprünge beschrieben werden können und mit dem andererseits schiefe und unsymmetrische Verteilungen der Returns motiviert werden. Es gibt mittlerweile zahlreiche Arbeiten, die sich mit Jump-Diffusions-Prozessen zur Beschreibung von Finanzzeitreihen, zur Bestimmung von Optionspreisen und zur Schätzung von Modellparametern beschäftigen (z.B. [4], [13]). Jump-Diffusions-Prozesse lassen sich nicht nur zur Beschreibung von Aktienkursen, sondern beispielsweise auch zur Modellierung von Kreditrisiken ([74]) oder zur Beschreibung kurzfristiger Zinsentwicklungen² ([1], [46], [48]) nutzen.

In der Literatur werden noch eine Vielzahl anderer Modelle zur möglichst adäquaten Beschreibung empirisch beobachteter Eigenschaften der Aktienkurse betrachtet und untersucht. So können schief und unsymmetrisch verteilte Returns auch durch Kursmodelle mit stochastischer Volatilität erklärt werden. Häufig wird die Volatilität als Mean-Reversion-Prozess angesehen. Während im Modell von Hull und White ([42]) die stochastische Volatilität unabhängig vom Wiener-Prozess ist, können im Heston-Modell (siehe [35], [72]) beide Prozesse voneinander abhängig sein. Um den Aktienkurs

¹Als Beispiel für solche Informationen sei das Bekanntwerden von Fusionsvorhaben, die Bekanntgabe wichtiger Wirtschaftsindikatoren mit unerwartetem Ausmaß oder die Bekanntgabe unerwarteter Zinsentscheidungen genannt.

²Üblicherweise wird den Mean-Reversion-Prozessen noch ein Sprungteil zugefügt.

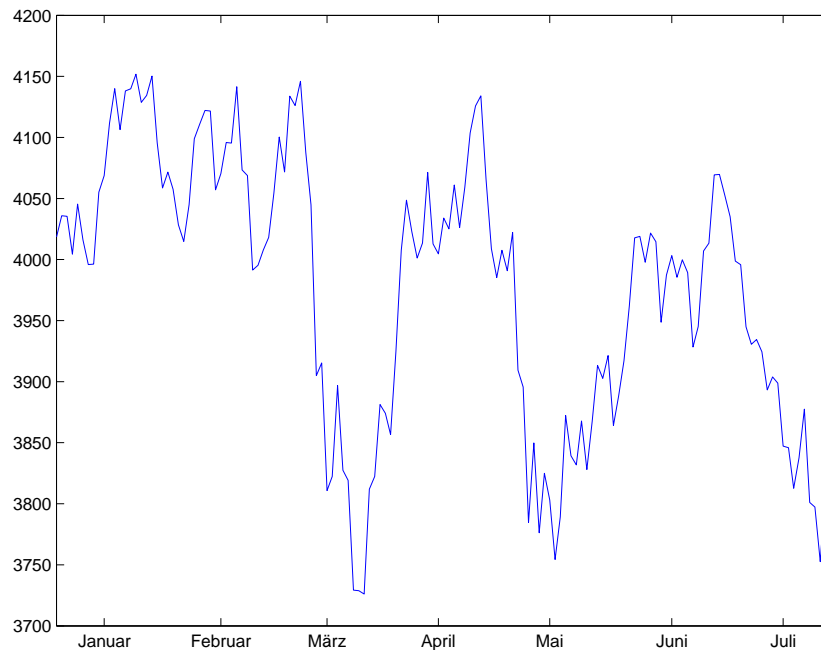


Abbildung 2.2: DAX Schlusskurse (Februar - Juli 2004)

möglichst realistisch nachzubilden, können die verschiedenen Bausteine einzelner Modelle natürlich noch vermischt werden (z.B. [15], [24], [25]). Das Problem der Parameterschätzung bzw. Modellkalibrierung wird in den meisten Fällen dadurch allerdings nicht vereinfacht.

Ziel dieser Arbeit ist es, die Möglichkeiten der Parameterschätzung für Jump-Diffusions-Prozesse näher zu untersuchen. Die Parameter, die das konkrete Verhalten von Jump-Diffusions-Prozessen spezifizieren, sollen möglichst stabil aus Marktdaten bestimmt werden. In praktischen Fällen hat man zur Parameterschätzung leider nicht so viele Daten³, wie man sie aus statistischer Sicht benötigen würde, um zuverlässige und stabile Werte für die Parameter zu erhalten. Daher wollen wir in dieser Arbeit die statistischen Zugänge zur Parameterschätzung mit Techniken aus der Theorie inverser Probleme kombinieren. Bei der Betrachtung von direkten Problemen sind die Parameter des stochastischen Modells zur Beschreibung der Aktienkurse gegeben und es ist beispielsweise eine Analyse des daraus resultierenden Kursverhaltens gesucht, oder aber es sollen Preise für derivative Produkte (Options- oder Forwardpreise) bestimmt werden. Bei direkten Problemen werden verschiedene Ursachen (Parameter) in ihre Wirkung (Preistrajektorien, Optionspreise) abgebildet. Bei den inversen Problemen (siehe z.B. [70], [7], [54], [32], [50], [38], [64]) hingegen versucht man aus einer festgestellten Wirkung, in unserem Fall die beobachteten Aktienkurse, auf die Ursache zu schließen, hier also die Parameter zu identifizieren. Viele inverse Probleme haben die Eigenschaft der Inkorrektheit (Schlechtgestellttheit), d.h. bereits kleine Messfehler in den Daten⁴ können

³Hat man beispielsweise Schlusskurse einer Aktie von einem Jahr, stehen ca. 250 Daten zur Verfügung.

⁴Bei Aktienkursen können Fehler einerseits bei der Notierung und Datenübermittlung und andererseits durch Beeinflussung der Preisbildung durch Steuern und Transaktionskosten auftreten.

zu starken Abweichungen in der Lösung führen. Um diese Inkorrektheit zu überwinden, regularisiert man das inverse Problem, d.h. man löst anstatt der instabilen Aufgabe ein benachbartes stabiles Ersatzproblem. Besonders ausgeprägt ist die Inkorrektheit bei inversen Problemen, die in unendlichdimensionalen Funktionenräumen formuliert werden. Selbst für lineare Vorwärtsoperatoren, die kompakt oder wenigstens stark glättend sind, ist die inverse Aufgabe dann instabil. Im Falle der Parameterschätzung bei stochastischen Prozessen liegt zwar in der Regel ein endlichdimensionales Problem vor, dessen Vorwärtsoperator aber nichtlinear ist. Wir werden in dieser Arbeit das Phänomen der Inkorrektheit für die Parameterbestimmung in Jump-Diffusions-Prozessen überzeugend deutlich machen. Sehr verschiedene Parameterkonstellationen können zu ähnlichen Kursverläufen führen. Dieser durch den glättenden Vorwärtsoperator auftretende Informationsverlust kann aber durch geschickten Einsatz von Apriori-Informationen deutlich abgemildert werden. Diesem Ziel dienen die in der Arbeit entwickelten Techniken und Algorithmen.

2.1 Kursmodell

In dem hier verwendeten Modell einer Assetpreisentwicklung werden sowohl gewöhnliche Informationen, die „normale“ stetige Kursänderungen implizieren, als auch außergewöhnliche Informationen, die Kurssprünge auslösen, zugelassen. Die Kurssprünge werden mit einem zusammengesetzten Poisson-Prozess, der „normale“ stetige Kursverlauf mit einer geometrischen Brownschen Bewegung modelliert.

Definition 2.1.1 *Es sei $\mathbf{S} = (S_t, t \in [0, \infty))$ ein stochastischer Prozess auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, \mathbb{P})$, der bezüglich der Filtration $(\mathfrak{F}_t, t \in [0, \infty))$ adaptiert ist, und der folgenden stochastischen Differentialgleichung*

$$dS_t = S_t((\mu - \lambda\nu)dt + \sigma dW_t) + S_t^- dN_t^z \quad (2.1)$$

mit Anfangswert⁵ S_0 genügt. Dabei ist

- $\mathbf{W} = (W_t, t \in [0, \infty))$ ein Standard-Wiener-Prozess,
- $(T_i)_{i \geq 1}$ die zu $\mathbf{N} = (N_t, t \in [0, \infty))$ gehörende Folge von Sprungzeitpunkten, d.h. $\forall i \geq 1$ gilt $N_{T_j^-} := \lim_{t \rightarrow T_j^-} N_t = N_{T_j^+} - 1$ mit $N_{T_j^+} := \lim_{t \rightarrow T_j^+} N_t$,
- $(Y_j)_{j \geq 1}$ eine Folge unabhängiger identisch verteilter Zufallsgrößen mit Erwartungswert $\nu + 1$ und
- $\mathbf{N}^z = (N_t^z, t \in [0, \infty))$ ein zu dem Poisson-Prozess $\mathbf{N} = (N_t, t \in [0, \infty))$ gehörender zusammengesetzter Poisson-Prozess mit Sprungintensität λ und Sprungamplituden $(Y_j - 1)$, d.h. wenn ein Sprung stattfindet, gilt $dN_t^z = N_{T_j^+}^z - N_{T_j^-}^z = Y_j - 1$ und somit $S_{T_j^+} = S_{T_j^-} + S_{T_j^-} dN_t^z = S_{T_j^-} + S_{T_j^-} (Y_j - 1) = S_{T_j^-} Y_j$.

⁵Wir werden in dieser Arbeit den Anfangswert S_0 ausschließlich als deterministische Größe betrachten.

Die Prozesse \mathbf{W} und \mathbf{N} sowie die Folge $(Y_j)_{j \geq 1}$ sind dabei unabhängig. Da die erste Komponente in (2.1) ein Diffusionsterm und der zweite Term eine Sprungkomponente ist, wird der Kurs \mathbf{S} auch als *Jump-Diffusions-Prozess* bezeichnet.

Die Verteilung der Sprunghöhen kann dabei beliebig gewählt werden. In [12] findet man einen Überblick über einige in der Literatur verwendete Modelle.

Bemerkung 2.1.2 Die stochastische Differentialgleichung (2.1) kann auch als Integralgleichung

$$\begin{aligned} S_t &= S_0 + \int_0^t S_u(\mu - \lambda\nu)du + \int_0^t S_u\sigma dW_u + \int_0^t S_u^- dN_u^z \\ &= S_0 + \int_0^t S_u(\mu - \lambda\nu)du + \int_0^t S_u\sigma dW_u + \sum_{j=1}^{N_t} S_{T_j^-}(Y_j - 1) \end{aligned} \quad (2.2)$$

formuliert werden. Dabei ist das stochastische Integral $\int_0^t S_u\sigma dW_u$ im Itô-Sinne definiert.

Für $N_t = 0$ vereinbaren wir $\sum_{j=1}^0 S_{T_j^-}(Y_j - 1) := 0$.

Annahme 2.1.3 Wir nehmen an, dass die Sprunghöhen Y_j logarithmisch normalverteilt mit Erwartungswert μ_Y und Varianz σ_Y^2 sind, d.h. $\ln Y_j \sim N(\mu_Y, \sigma_Y^2)$, so dass für den Erwartungswert der Sprungamplituden

$$\nu = \mathbb{E}(Y_j - 1) = \exp\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right) - 1 \quad (2.3)$$

gilt.

Ein solches Kursmodell wurde bereits in den 70er-Jahren von Merton in dem Artikel [55] „Option Pricing when underlying stock returns are discontinuous“ (vgl. auch [56, Chapter 9]) verwendet.

Insgesamt beschreiben somit fünf Parameter das konkrete Kursverhalten unseres Jump-Diffusions-Prozesses. Die Parameter μ und σ bestimmen dabei Drift und Volatilität des Diffusionsterms, λ legt die Sprunghäufigkeit fest und μ_Y und σ_Y spezifizieren die Sprünge. Wir fassen diese fünf Parameter im Parametervektor $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)$ zusammen. Der Fall $\sigma = 0$ nimmt dabei eine Sonderrolle ein, da der Diffusionsprozess in diesem Fall zu einer deterministischen linearen Funktion mit Anstieg μ entartet. Im Weiteren betrachten wir daher nur Parameter aus dem Definitionsbereich

$$D = \{\underline{p} \in \mathbb{R}^5 : \sigma > 0, \lambda \geq 0, \sigma_Y \geq 0\}. \quad (2.4)$$

Später werden wir den Fall $\sigma = 0$ bzw. den Grenzfall $\sigma \rightarrow 0$ noch genauer untersuchen.

An dieser Stelle können wir kurz das Vorwärtsproblem betrachten. Wir bezeichnen für jeweils fixiertes $\omega \in \Omega$ die entsprechende Trajektorie $(S_t(\omega), t \in [0, \infty))$ des

Prozesses $\mathbf{S} = (S_t, t \in [0, \infty))$ mit $S(\omega)$ bzw. S . Eine konkrete Realisierung werden wir zur besseren Unterscheidung mit S^* bezeichnen. Somit lässt sich das allgemeine Vorwärtsproblem wie folgt beschreiben. Jedem Parametervektor \underline{p} wird ein stochastischer Prozess zugeordnet. Die Eigenschaften dieses Jump-Diffusions-Prozesses hängen natürlich von dem jeweiligen Parametervektor ab. Für eine konkrete Zufallssituation erhalten wir dann eine Trajektorie des stochastischen Prozesses. Die folgende Abbildung verdeutlicht nochmals das allgemeine Vorwärtsproblem.

$$\boxed{\underline{p} \in D \mapsto \mathbf{S} = \mathbf{S}(\underline{p}) : \omega \in \Omega \mapsto S = S(\underline{p}, \omega)}$$

Zur Bestimmung der Parameter aus Kursdaten steht in praxi genau eine konkrete Trajektorie S^* zur Verfügung. Leider kennen wir die Werte dieser Trajektorie nicht für jeden Zeitpunkt $t \in [0, \infty)$, sondern nur zu Zeitpunkten $t \in T$. Dabei ist T eine echte Teilmenge der nichtnegativen reellen Zahlen, d.h. $T \subsetneq [0, \infty)$. In praktischen Anwendungen sind im Allgemeinen die Werte von S^* nur zu diskreten Zeitpunkten über einem endlichen Zeithorizont gegeben, so dass $T = \{t_0, t_1, \dots, t_n\}$ gilt. Üblicherweise stehen historische Kurse einer Aktie oder eines Aktienindizes zur Verfügung, um die Parameter zu bestimmen. Denkt man dabei beispielsweise an Tagesschlusskurse, hat man die Daten automatisch über einem äquidistanten Zeitgitter t_0, t_1, \dots, t_n gegeben, da $\tau(i) = t_i - t_{i-1} = \frac{1}{250} = \tau$ für alle $i = 1, \dots, n$ gilt. Allerdings existieren auch Kursnotierungen zu unregelmäßigen Zeiten, so dass in diesen Fällen $\tau(i)$ wirklich von i abhängt. Dieser Fall tritt zum Beispiel bei der Betrachtung von Intradaykursen auf. Es gibt aber Techniken, wie man aus solchen unregelmäßigen Notierungen Daten in äquidistanten Zeitpunkten gewinnen kann. Eine Möglichkeit besteht beispielsweise darin, zunächst durch Interpolation der Daten eine bezüglich der Zeit stetige Trajektorie zu berechnen. Anschließend tastet man dann die Kurse auf dem gewünschten äquidistanten Zeitgitter ab (vgl. [67]). Bei dieser Methode sollte man aber darauf achten, dass durch die Interpolation der Kurse keine künstlichen Abhängigkeiten erzeugt werden. Im Folgenden werden wir davon ausgehen, dass die Daten bezüglich eines äquidistanten Zeitgitters gegeben sind.

Annahme 2.1.4 Die Kursdaten $S^* = (S_{t_0}^*, \dots, S_{t_n}^*)$, aus denen der Parameter $\underline{p}^* \in D$ identifiziert werden soll, sind über dem äquidistanten Zeitgitter t_0, t_1, \dots, t_n mit Zeitabstand $\tau = t_i - t_{i-1}$ ($i = 1, \dots, n$) gegeben.

Werden zur Bestimmung des Parameters empirische Kursdaten genutzt, schreiben wir $S^\delta = (S_{t_0}^\delta, \dots, S_{t_n}^\delta)$ anstelle von $S^* = (S_{t_0}^*, \dots, S_{t_n}^*)$, da diese Daten häufig wie bereits oben diskutiert mit einem Datenfehler δ behaftet sind. Zur besseren Unterscheidung werden wir in dieser Arbeit exakte Daten mit \star und empirische Daten mit δ markieren.

Bei dem allgemeinen inversen Problem versucht man nun aus der Trajektorie S^* den zugehörigen Parametervektor $\underline{p}^* \in D$ zu identifizieren. Natürlich ist dabei die Zufallssituation $\omega \in \Omega$, die diese Trajektorie generiert hat, unbekannt. Ferner können noch andere Parametervektoren existieren, die mit einer entsprechenden Zufallssituation Trajektorien generieren, die für alle Zeitpunkte $t \in T$ mit S^* übereinstimmen. An dieser Stelle wird der Unterschied zwischen unserem stochastischen Problem und einem rein

deterministischen Problem besonders deutlich. Die Transformation eines Parametervektors auf eine Trajektorie hat zufallsbedingt einen gewissen glättenden Charakter, d.h. verschiedene Ursachen (Parametern) können in eine ähnliche Wirkung (Trajektorie) abgebildet werden. Dabei müssen wir die folgenden Situationen unterscheiden. Einerseits können verschiedene Parametervektoren in Prozesse mit ähnlicher Wahrscheinlichkeitsverteilung abgebildet werden. In diesem Fall werden sich auch viele Trajektorien beider Prozesse ähneln. Andererseits ist es aber auch möglich, dass sich zufällig einzelne Trajektorien völlig verschiedener Jump-Diffusions-Prozesse ähneln können. Daher muss bei der Lösung des inversen Problems die Verteilung des Prozesses berücksichtigt werden. Aus diesem Grunde bestimmen wir zunächst die Lösung der stochastischen Differentialgleichung (2.1), um dann nähere Aussagen über die Verteilung treffen zu können.

Für das Kursmodell der Jump-Diffusions-Prozesse aus Definition 2.1.1 bilden ein Wiener-Prozess und ein Poisson-Prozess die stochastischen Grundbausteine. Diese beiden Prozesse gehören zur Klasse der Lévy-Prozesse.

Definition 2.1.5 *Ein stochastischer Prozess $\mathbf{X} = (X_t, t \in [0, \infty))$ mit $X_0 = 0$ auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, \mathbb{P})$, der bezüglich der Filtration $(\mathfrak{F}_t, t \in [0, \infty))$ adaptiert ist, heißt Lévy-Prozess, falls er von der Vergangenheit unabhängige und stationäre Zuwächse besitzt und stetig in Wahrscheinlichkeit ist, d.h.*

1. $X_t - X_s$ ist für alle $0 \leq s < t < \infty$ von \mathfrak{F}_s unabhängig,
2. $X_t - X_s$ besitzt für alle $0 \leq s < t < \infty$ die gleiche Verteilung wie X_{t-s} und
3. $X_t \xrightarrow[t \rightarrow s]{\mathbb{P}} X_s$.

Bemerkung 2.1.6 Man kann zeigen, dass jeder Lévy-Prozess eine eindeutige Modifikation mit rechtsseitig stetigen Trajektorien und linksseitig existierenden Grenzwerten besitzt. Diese Eigenschaft wird auch als càdlàg („continu à droite, limite à gauche“ – vgl. [12]) bezeichnet.

Nach Theorem 9 und dessen Folgerung in [63, Chapter 2] gehören die Lévy-Prozesse wiederum zur größeren Klasse der Semimartingale⁶. Wir werden nun zeigen, dass der Jump-Diffusions-Prozess mit der Integraldarstellung (2.2) zur Klasse der Semimartingale zählt.

Die stochastischen Integrale $\int_0^t S_u \sigma dW_u$ und $\int_0^t S_u - dN_u^z$ in (2.2) sind nach Theorem 19 in [63] wieder Semimartingale, da sowohl der Wiener-Prozess als auch der Poisson-Prozess Lévy-Prozesse und somit Semimartingale sind. Folglich ist der Prozess \mathbf{S} auch ein Semimartingal, da die Summe von Semimartingalen wieder ein Semimartingal ergibt und das Integral $\int_0^t S_u (\mu - \lambda \nu) du$ mit deterministischen Integrator trivialerweise ein Semimartingal ist. Damit zählen also die Jump-Diffusions-Prozesse zur Klasse

⁶Zur Definition von Semimartingalen siehe [63, Chapter 2].

der Semimartingale, für die ein verallgemeinertes Itô-Kalkül existiert. Mit deren Hilfe können wir nun die Lösung der stochastischen Differentialgleichung (2.1) angeben. Hierfür benötigen wir allerdings einige Begriffe und Aussagen.

Definition 2.1.7 *Es sei $\mathbf{X} = (X_t, t \in [0, \infty))$ ein Semimartingal. Der stochastische Prozess $[\mathbf{X}, \mathbf{X}] = ([X, X]_t, t \in [0, \infty))$ mit⁷*

$$[X, X]_t = X_t^2 - 2 \int_0^t X_{s-} dX_s$$

heißt Prozess der quadratischen Variation von \mathbf{X} .

Satz 2.1.8 (vgl. [63]) *Für den Prozess der quadratischen Variation $[\mathbf{X}, \mathbf{X}]$ von \mathbf{X} gilt mit $\Delta X = X_t - X_{t-}$ gerade*

1. $[X, X]_0 = X_0^2$ und
2. $\Delta[X, X]_t = (\Delta X)_t^2$.

Beweis. Es gelten die folgenden Gleichungen

$$\begin{aligned} (\Delta X)_t^2 &= (X_t - X_{t-})^2 = X_t^2 - 2X_t X_{t-} + X_{t-}^2 \\ &= X_t^2 - X_{t-}^2 + 2X_{t-}(X_{t-} - X_t) \\ &= \Delta(X^2)_t - 2X_{t-}(\Delta X_t) \\ &= \Delta[X, X]_t \end{aligned}$$

□

Mit diesen Begriffen kann nun die Itô-Formel für Semimartingale beschrieben werden, deren Beweis in [63] nachzulesen ist.

Satz 2.1.9 *Es sei $(X_t, t \in [0, \infty))$ ein Semimartingal und f eine zweimal stetig differenzierbare reelle Funktion. Dann ist $(f(X_t), t \in [0, \infty))$ ebenfalls ein Semimartingal und es gilt*

$$\begin{aligned} f(X_t) &= f(X_0) + \int_{0+}^t f'(X_{u-}) dX_u + \frac{1}{2} \int_{0+}^t f''(X_{u-}) d[X, X]_u^c \\ &\quad + \sum_{0 < u \leq t} (f(X_u) - f(X_{u-}) - f'(X_{u-}) \Delta X_u), \end{aligned} \tag{2.5}$$

wobei $[\mathbf{X}, \mathbf{X}]^c$ den pfadweise stetigen Anteil des Prozesses der quadratischen Variation $[\mathbf{X}, \mathbf{X}]$ bezeichnet.

⁷Wir definieren dabei $X_{0-} := 0$.

Für Semimartingale mit stetigen Pfaden vereinfacht sich Formel (2.5) gerade zu

$$f(X_t) = f(X_0) + \int_0^t f'(X_u) dX_u + \frac{1}{2} \int_0^t f''(X_u) d[X, X]_u,$$

welche der Itô-Formel für Itô-Prozesse entspricht (siehe z.B. [45], [61]). Zur Vereinfachung setzen wir im Folgenden

$$\tilde{\mu} := \mu - \lambda\nu - \frac{1}{2}\sigma^2. \quad (2.6)$$

Theorem 2.1.10 *Der Prozess*

$$S_t = S_0 \exp(\tilde{\mu}t + \sigma W_t) \prod_{j=1}^{N_t} Y_j, \quad (2.7)$$

mit $\prod_{j=k}^l Y_j := 1$ für $l < k$ ist Lösung der stochastischen Differentialgleichung (2.1) aus Definition 2.1.1, d.h. (2.7) ist ein Jump-Diffusions-Prozess.

Beweis. Für den Prozess (2.1) gilt nach dem zweiten Punkt in Satz 2.1.8 für die quadratische Variation nach Grenzwertbildung

$$\begin{aligned} d[S, S]_u &= (dS)_u^2 = S_u^2 ((\mu - \lambda\nu)^2 (du)^2 + 2(\mu - \lambda\nu)\sigma du dW_u + \sigma^2 (dW_u)^2) \\ &\quad + 2S_u S_{u-} ((\mu - \lambda\nu) du + \sigma dW_u) dN_u^z + S_{u-}^2 (dN_u^z)^2 \end{aligned}$$

und wegen $(du)^2 = 0$, $du dW_u = 0$ und $(dW_u)^2 = du$ folglich

$$d[S, S]_u = S_u^2 \sigma^2 du + dN_u^z (2S_u S_{u-} ((\mu - \lambda\nu) du + \sigma dW_u) + S_{u-}^2 (dN_u^z)).$$

Somit erhalten wir für den pfadweise stetigen Anteil der quadratischen Variation $d[S, S]_u^c = \sigma^2 S_u^2 du$, da der pfadweise stetige Anteil des zweiten Summanden verschwindet. Mit Satz 2.1.9 können wir zu den logarithmierten Kursen übergehen. Im Falle eines Sprunges gilt $\Delta S_u = S_{u+} - S_{u-} = S_{u-} Y_j$ sowie $\frac{S_{T_j^+}}{S_{T_j^-}} = Y_j$ (vgl. Def. 2.1.1) und somit erhalten wir mit (2.6)

$$\begin{aligned} \ln S_t &= \ln S_0 + \int_{0+}^t \frac{1}{S_{u-}} dS_u - \frac{1}{2} \int_{0+}^t \frac{1}{S_{u-}^2} d[S, S]_u^c + \sum_{0 < u \leq t} (\ln S_u - \ln S_{u-} - \frac{1}{S_{u-}} \Delta S_u) \\ &= \ln S_0 + \int_0^t \frac{S_u}{S_{u-}} (\mu - \lambda\nu) du + \int_0^t \frac{S_u}{S_{u-}} \sigma dW_u + \int_{0+}^t \frac{S_{u-}}{S_{u-}} dN_u^z - \frac{1}{2} \int_0^t \frac{S_u^2}{S_{u-}^2} \sigma^2 du \\ &\quad + \sum_{j=1}^{N_t} (\ln \frac{S_{T_j}}{S_{T_j^-}} - \frac{1}{S_{T_j^-}} S_{T_j^-} (Y_j - 1)) \quad (2.8) \\ &= \ln S_0 + \int_0^t (\mu - \lambda\nu - \frac{\sigma^2}{2}) du + \int_0^t \sigma dW_u + \sum_{j=1}^{N_t} (Y_j - 1) + \sum_{j=1}^{N_t} (\ln Y_j - (Y_j - 1)) \\ &= \ln S_0 + \left(\mu - \lambda\nu - \frac{\sigma^2}{2} \right) t + \sigma W_t + \sum_{j=1}^{N_t} \ln Y_j \end{aligned}$$

und daraus schließlich Beziehung (2.7) □

Gleichung (2.8) hat in Differentialschreibweise die Darstellung (vgl. [74])

$$d \ln S_t = \tilde{\mu} dt + \sigma dW_t + dN_t^{z^*} ,$$

wobei $(N_t^{z^*}, t \in [0, \infty))$ einen zusammengesetzten Poisson-Prozess mit Sprunghöhen $\ln Y_j$ bezeichnet.

In der Finanzmathematik ist es üblich, die Returns eines Assets zu betrachten, da diese oftmals stationär und unabhängig sind. Für die Kurse des Assets trifft diese Eigenschaft in den meisten Fällen nicht zu, da man beispielsweise für Aktienkurse im Allgemeinen langfristig einen positiven Trend erwartet. Falls die Returns stationär und unabhängig sind, kann man aus den Kursen $S^\delta = (S_{t_0}^\delta, \dots, S_{t_n}^\delta)$ die Returns berechnen, und erhält dann n unabhängige Realisierungen einer Zufallsgröße, anstatt jeweils eine Realisierung zu jedem Zeitpunkt t_i . Der Übergang von den Kursen $S^\delta = (S_{t_0}^\delta, \dots, S_{t_n}^\delta)$ zu den Returns mit Zeitabstand τ

$$r^\delta = (r_1^\delta, \dots, r_n^\delta) \quad \text{mit} \quad r_i^\delta = \ln \frac{S_{t_i}^\delta}{S_{t_{i-1}}^\delta}, \quad i = 1, \dots, n \quad (2.9)$$

ist allerdings nur von Vorteil, falls deren Eigenschaften bekannt sind. Daher werden wir im folgenden Abschnitt die Returns etwas näher untersuchen.

2.2 Eigenschaften der Returns

2.2.1 Stationarität der Returns

Definition 2.2.1 *Der Return eines Finanzgutes, beginnend zum Zeitpunkt x mit Zeitabstand (Lag) τ , ist wie folgt definiert*

$$r_{x, x+\tau} = \ln \frac{S_{x+\tau}}{S_x} .$$

Mit Hilfe von Satz 2.1.10 können wir den stochastischen Prozess der Returns eines Jump-Diffusions-Prozesses sofort angeben. Die Returns $r_{0, \tau} = \ln \frac{S_\tau}{S_0}$ haben dann die Darstellung

$$r_{0, \tau} = \tilde{\mu} \tau + \sigma W_\tau + \sum_{j=1}^{N_\tau} \ln Y_j . \quad (2.10)$$

Wenn wir nun τ für Werte aus $[0, \infty)$ betrachten, erhalten wir den stochastischen Prozess $\mathbf{r} = (r_{0, \tau}, \tau \in [0, \infty))$ der Returns eines Jump-Diffusions-Prozesses.

Lemma 2.2.2 *Der Prozess \mathbf{r} der Returns mit der Darstellung (2.10) ist ein Lévy-Prozess.*

Beweis. Die Adaptiertheit sowie die Eigenschaft der stochastischen Stetigkeit des Prozesses \mathbf{r} folgen unmittelbar aus der Adaptiertheit und stochastischen Stetigkeit des Preisprozesses \mathbf{S} . Für die Zuwächse der Returns gilt

$$\begin{aligned} \Delta r(\tau, \Delta\tau) &:= r_{0,\tau+\Delta\tau} - r_{0,\tau} = \tilde{\mu}\Delta\tau + \sigma(W_{\tau+\Delta\tau} - W_\tau) + \sum_{j=N_\tau+1}^{N_{\tau+\Delta\tau}} \ln Y_j \\ &\stackrel{(d)}{=} \tilde{\mu}\Delta\tau + \sigma(W_{\tau+\Delta\tau} - W_\tau) + \sum_{j=1}^{N_{\tau+\Delta\tau}-N_\tau} \ln Y_j \\ &\stackrel{(d)}{=} \tilde{\mu}\Delta\tau + \sigma W_{\Delta\tau} + \sum_{j=1}^{N_{\Delta\tau}} \ln Y_j = r_{0,\Delta\tau}, \end{aligned} \quad (2.11)$$

weil die Sprünge $\ln Y_j$ für alle j unabhängig und identisch verteilt sind und der Wiener-Prozess \mathbf{W} und der Poisson-Prozess \mathbf{N} Lévy-Prozesse sind, also stationäre Zuwächse besitzen. Da die Zuwächse der Returns $\Delta r(\tau, \Delta\tau)$ nur von $\Delta\tau$, nicht aber von τ abhängen, sind diese stationär. Die Unabhängigkeit der Zuwächse folgt unmittelbar aus der Unabhängigkeit der Zuwächse des Wiener-Prozesses und des Poisson-Prozesses sowie der Unabhängigkeit von \mathbf{W} , \mathbf{N} und den Sprüngen $(\ln Y_j)_{j \geq 1}$ untereinander. Somit ist der Prozess \mathbf{r} ein Lévy-Prozess \square

Bemerkung 2.2.3 Da der deterministische Startwert S_0 bekannt ist, können wir ohne Informationsverlust von den Aktienkursen zu den Returns übergehen.

Ebenfalls interessant ist der Fall, dass wir das Lag τ der Returns fixieren, und den Anfangszeitpunkt x variieren, d.h. wir betrachten nun $r_{x,x+\tau}$ für $x \in [0, \infty)$.

Lemma 2.2.4

- (i) Die Returns $r_{x,x+\tau}$ und $r_{0,\tau}$ besitzen für alle $x \geq 0$ die gleiche Verteilung, d.h. für fixiertes Lag τ sind die Returns bezüglich des Startzeitpunktes stationär.
- (ii) Die Returns $r_{x,x+\tau}$ und $r_{y,y+\tau}$ sind für alle x und y mit $|y - x| \geq \tau$ unabhängig.

Beweis. Mit (2.11) gilt für den Return $r_{x,x+\tau}$

$$\begin{aligned} r_{x,x+\tau} &= \ln \frac{S_{x+\tau}}{S_x} = \ln S_{x+\tau} - \ln S_0 + \ln S_0 - \ln S_x \\ &= r_{0,x+\tau} - r_{0,x} = \Delta r(x, \tau) \stackrel{(d)}{=} \tilde{\mu}\tau + W_\tau + \sum_{j=1}^{N_\tau} \ln Y_j = r_{0,\tau}, \end{aligned}$$

und somit Behauptung (i). Die zweite Behauptung folgt unmittelbar aus der Unabhängigkeit der Zuwächse \square

Da der Return $r_{x,x+\tau}$ die gleiche Verteilung wie $r_{0,\tau}$ besitzt und der Startzeitpunkt x keine wesentliche Rolle spielt, schreiben wir im Folgenden für die Returns mit Lag τ nur

r_τ . Aus (2.11) erhalten wir für die Zuwächse der Returns die Beziehung $\Delta r(\tau, \Delta\tau) \stackrel{(d)}{=} r_{0,\Delta\tau} = r_{\Delta\tau}$, d.h. die Zuwächse haben die gleiche Verteilung wie Returns mit Lag $\Delta\tau$.

Die zu den Zeitpunkten t_0, t_1, \dots, t_n bekannte Trajektorie $S^* = (S_{t_0}^*, \dots, S_{t_n}^*)$ ist eine konkrete Realisierung des Zufallsvektors $S(\omega) = (S_{t_0}(\omega), S_{t_1}(\omega), \dots, S_{t_n}(\omega))$, dessen Komponenten alle unterschiedlich verteilt sind. Der Übergang von den Kursdaten $S^* = (S_{t_0}^*, \dots, S_{t_n}^*)$ zu den Returns $r^* = (r_1^*, r_2^*, \dots, r_n^*)$ mit

$$r_i^* := \ln \frac{S_{t_i}^*}{S_{t_{i-1}}^*} \quad (2.12)$$

liefert uns zunächst eine Realisierung des Zufallsvektors

$$r(\omega) = (r_{t_0, t_1}(\omega), r_{t_1, t_2}(\omega), \dots, r_{t_{n-1}, t_n}(\omega)) \quad (2.13)$$

und wegen der Stationarität ($r_{t_{i-1}, t_i} = r_{t_{i-1}, t_{i-1}+\tau} \stackrel{(d)}{=} r_\tau$) und Unabhängigkeit der Returns Realisierungen von n unabhängigen wie r_τ verteilten Zufallsgrößen. Im Weiteren untersuchen wir nun die Verteilung sowie die Momente der Returns r_τ näher.

2.2.2 Verteilung der Returns

Für die Returns mit Lag τ gilt nach (2.10) gerade $r_\tau = \tilde{\mu}\tau + \sigma W_\tau + \sum_{j=1}^{N_\tau} \ln Y_j$. Fixieren wir die Anzahl der Sprünge auf j , so ist der Return r_τ normalverteilt mit Erwartungswert $\tilde{\mu}\tau + j\mu_Y$ und Varianz $\sigma^2\tau + j\sigma_Y^2$. An dieser Stelle wird die Bedeutung von Annahme 2.1.3 deutlich, da wir die Normalverteilung von $\ln Y_j$ benötigen. Mit dem Satz der totalen Wahrscheinlichkeit erhalten wir nun für jeden Parametervektor $\underline{p} \in D$ die Verteilungsfunktion der Returns r_τ

$$\begin{aligned} F_{r_\tau}(x, \underline{p}) &= \mathbb{P}(r_\tau \leq x) = \sum_{j=0}^{\infty} \mathbb{P}(N_\tau = j) \mathbb{P}(r_\tau \leq x \mid N_\tau = j) \\ &= \sum_{j=0}^{\infty} \frac{e^{(-\lambda\tau)} (\lambda\tau)^j}{j!} \Phi \left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \right), \end{aligned} \quad (2.14)$$

wobei $\Phi(x)$ die Verteilungsfunktion der Standardnormalverteilung bezeichnet. Ableitung der Verteilungsfunktion (2.14) nach x für festen Parametervektor $\underline{p} \in D$ und festes Lag τ liefert uns schließlich die folgende Darstellung der Dichtefunktion

$$f_{r_\tau}(x, \underline{p}) = F'_{r_\tau}(x, \underline{p}) = \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi \left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \right). \quad (2.15)$$

Dabei bezeichnet $\phi(x)$ die Dichtefunktion einer standardisierten normalverteilten Zufallsgröße, d.h.

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2} \right).$$

Man überprüft leicht, dass für jeden festen Parametervektor $\underline{p} \in D$ und festes Lag τ die Funktion (2.14) wirklich eine Verteilungsfunktion und (2.15) eine Dichtefunktion ist. Außerdem sehen wir anhand der Verteilungs- bzw. Dichtefunktion der Returns, dass diese eine unendliche Überlagerung gewichteter normalverteilter Zufallsgrößen sind, was den Hintergrund der Inkorrektheit des inversen Problems bildet. Wir kennen nun die Verteilung der Returns. Viele statistische Eigenschaften kann man aber besser erkennen, indem man die Momente und Semiinvarianten betrachtet. Die statistischen Kenngrößen Schiefe und Exzess kann man beispielsweise direkt von den Momenten bzw. Semiinvarianten ableiten. Wir berechnen die Momente und Semiinvarianten mit Hilfe der charakteristischen Funktion.

Lemma 2.2.5 *Die charakteristische Funktion der Returns r_τ mit Parametervektor \underline{p} besitzt die Darstellung*

$$\varphi_{r_\tau}(\theta, \underline{p}) = \exp \left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\tau\theta^2 + \lambda\tau \left(\exp \left(-\frac{\sigma_Y^2}{2}\theta^2 + i\mu_Y\theta \right) - 1 \right) \right). \quad (2.16)$$

Beweis. Die in (2.17) dargestellten Gleichungen liefern Formel (2.16)

$$\begin{aligned} \varphi_{r_\tau}(\theta, \underline{p}) &= \mathbb{E} \exp(ir_\tau\theta) = \mathbb{E} \exp \left(i\theta(\tilde{\mu}\tau + \sigma W_\tau) + i\theta \sum_{j=1}^{N_\tau} \ln Y_j \right) \\ &= \mathbb{E} \exp(i\theta(\tilde{\mu}\tau + \sigma W_\tau)) \mathbb{E} \exp \left(i\theta \sum_{j=1}^{N_\tau} \ln Y_j \right) \\ &= \exp \left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\theta^2\tau \right) \sum_{k=0}^{\infty} \mathbb{E} \left(\exp \left(i\theta \sum_{j=1}^{N_\tau} \ln Y_j \right) \middle| N_\tau = k \right) \mathbb{P}(N_\tau = k) \\ &= \exp \left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\theta^2\tau \right) \sum_{k=0}^{\infty} (\mathbb{E}(\exp(i\theta \ln Y_1)))^k \exp(-\lambda\tau) \frac{(\lambda\tau)^k}{k!} \quad (2.17) \\ &= \exp \left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\theta^2\tau \right) \sum_{k=0}^{\infty} (\varphi_{\ln Y_1}(\theta))^k \exp(-\lambda\tau) \frac{(\lambda\tau)^k}{k!} \\ &= \exp \left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\theta^2\tau + \lambda\tau(\varphi_{\ln Y_1}(\theta) - 1) \right) \\ &= \exp \left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\theta^2\tau + \lambda\tau \left(\exp \left(i\mu_Y\theta - \frac{\sigma_Y^2}{2}\theta^2 \right) - 1 \right) \right) \end{aligned}$$

□

Bemerkung 2.2.6 Im obigen Beweis zeigt sich deutlich, dass Annahme 2.1.3 zur Berechnung der charakteristischen Funktion nur eine untergeordnete Rolle spielt. Wir können die charakteristische Funktion der Returns r_τ eines Jump-Diffusions-Prozesses für eine beliebige Verteilungsdichte f_J der Sprünge $\ln Y_j$ berechnen, wenn deren charakteristische Funktion $\varphi_{\ln Y_1}(\theta)$ bekannt ist. Im Gegensatz dazu benötigen wir zur Berechnung der Verteilungsfunktion normalverteilte Sprünge $\ln Y_j$. Deshalb ist für viele

Jump-Diffusions-Prozesse, die eine andere Sprungverteilung besitzen, Verteilungsfunktion bzw. Verteilungsdichte nicht direkt bekannt. Man kann diese aber durch eine inverse Fouriertransformation der charakteristischen Funktion berechnen. Nach Lemma 2.2.2 ist der Prozess \mathbf{r} ein Lévy-Prozess. Deshalb hätten wir die charakteristische Funktion auch mit Hilfe der Lévy-Khinchin-Formel (vgl. [67]) herleiten können. Für den einfachen Fall, dass die Sprünge durch einen zusammengesetzten Poisson-Prozess modelliert werden, hat diese die Darstellung

$$\varphi_{r_\tau}(\theta, \underline{p}) = \exp \left(\tau \left(i\tilde{\mu}\theta - \frac{\sigma^2}{2}\theta^2 + \int_{-\infty}^{\infty} (\exp(i\theta x) - 1) \nu(dx) dx \right) \right),$$

die wir wegen⁸ $\nu(dx) = \lambda f_J(x)dx$ und $\int_{-\infty}^{\infty} (\exp(i\theta x) - 1) f_J(x)dx = \varphi_{\ln Y_1}(\theta) - 1$ bereits im Beweis von Lemma 2.2.5 (vgl. (2.17)) hergeleitet haben.

Die charakteristische Funktion (2.16) ist eine Exponentialfunktion und daher beliebig oft differenzierbar.

Folgerung 2.2.7 *Die charakteristische Funktion $\varphi_{r_\tau}(\theta, \underline{p})$ der Returns r_τ ist bezüglich θ für alle $\underline{p} \in D$ unendlich oft differenzierbar, d.h. es gilt $\varphi_{r_\tau}(\cdot, \underline{p}) \in C^\infty(\mathbb{R})$ für alle $\underline{p} \in D$.*

Im Gegensatz zur Dichtefunktion (2.15), die einer unendlichen Reihenentwicklung genügt, ist die charakteristische Funktion von einfacher Form. Der allgemeine Zusammenhang zwischen der k -ten Ableitung der charakteristischen Funktion einer beliebigen Zufallsgröße ξ mit $\mathbb{E}|\xi|^n < \infty$ und den Momenten bzw. Semiinvarianten lautet

$$m_k = \mathbb{E}\xi^k = \frac{\varphi_\xi^{(k)}(0)}{i^k} \quad \text{und} \quad s_k = \frac{\ln \varphi_\xi^{(k)}(0)}{i^k} \quad \forall k \leq n. \quad (2.18)$$

Aus der zweiten Gleichung in (2.18) erhalten wir mit Hilfe der ersten Ableitungen der logarithmierten charakteristischen Funktion das folgende System für die Semiinvarianten

⁸Das Maß ν ist das so genannte Lévy-Maß, das die Bedingungen $\nu(0) = 0$, $\int \nu(dx) = \lambda$ und $\int (|x|^2 \wedge 1) \nu(dx) < \infty$ erfüllt. Im Falle des zusammengesetzten Poisson-Prozesses ist das Lévy-Maß immer endlich.

der Returns.

$$\begin{aligned}
s_{\tau,1}(\underline{p}) &= (\lambda\mu_Y + \tilde{\mu})\tau \\
s_{\tau,2}(\underline{p}) &= ((\sigma_Y^2 + \mu_Y^2)\lambda + \sigma^2)\tau \\
s_{\tau,3}(\underline{p}) &= (3\sigma_Y^2 + \mu_Y^2)\lambda\tau\mu_Y \\
s_{\tau,4}(\underline{p}) &= (3\sigma_Y^4 + 6\sigma_Y^2\mu_Y^2 + \mu_Y^4)\lambda\tau \\
s_{\tau,5}(\underline{p}) &= (15\sigma_Y^4 + 10\sigma_Y^2\mu_Y^2 + \mu_Y^4)\lambda\tau\mu_Y \\
s_{\tau,6}(\underline{p}) &= (15\sigma_Y^6 + 45\sigma_Y^4\mu_Y^2 + 15\sigma_Y^2\mu_Y^4 + \mu_Y^6)\lambda\tau \\
s_{\tau,7}(\underline{p}) &= (105\sigma_Y^6 + 105\sigma_Y^4\mu_Y^2 + 21\sigma_Y^2\mu_Y^4 + \mu_Y^6)\lambda\tau\mu_Y \\
s_{\tau,8}(\underline{p}) &= (105\sigma_Y^8 + 420\sigma_Y^6\mu_Y^2 + 210\sigma_Y^4\mu_Y^4 + 28\sigma_Y^2\mu_Y^6 + \mu_Y^8)\lambda\tau \\
&\vdots
\end{aligned} \tag{2.19}$$

Analog könnten wir auch mittels der ersten Gleichung in (2.18) die Momente berechnen. Allerdings ist dies aufgrund der Struktur der charakteristischen Funktion (2.16) umständlich. Wir werden zur Berechnung der Momente den allgemeinen Zusammenhang zwischen Momenten und Semiinvarianten nutzen, der in dem folgenden Satz (vgl. [66]) dargestellt ist.

Satz 2.2.8 *Es sei ξ eine Zufallsgröße mit $E|\xi|^n < \infty$. Dann gilt $\forall k \leq n$*

$$m_k^\xi = \sum_{\lambda^{(1)} + \dots + \lambda^{(q)} = k} \frac{1}{q!} \frac{k!}{\lambda^{(1)}! \dots \lambda^{(q)}!} \prod_{p=1}^q s_{\lambda^{(p)}}^\xi$$

und

$$s_k^\xi = \sum_{\lambda^{(1)} + \dots + \lambda^{(q)} = k} \frac{(-1)^{q-1}}{q} \frac{k!}{\lambda^{(1)}! \dots \lambda^{(q)}!} \prod_{p=1}^q m_{\lambda^{(p)}}^\xi, \tag{2.20}$$

wobei $\sum_{\lambda^{(1)} + \dots + \lambda^{(q)} = k}$ die Addition über alle geordneten Mengen von $\lambda^{(p)} \in \mathbb{N}$ mit $\sum_{p=1}^q \lambda^{(p)} = k$ darstellt.

Aus der Existenz der $2n$ -ten Ableitung der charakteristischen Funktion im Punkt $\theta = 0$ folgt die Existenz der Momente bis zur $2n$ -ten Ordnung. Wegen 2.2.8 können wir dann auch auf die Existenz der Semiinvarianten bis zur $2n$ -ten Ordnung schlussfolgern. Da uns Folgerung 2.2.7 für jedes beliebige $n \in \mathbb{N}$ die Existenz der $2n$ -ten Ableitung der charakteristischen Funktion im Punkt $\theta = 0$ sichert, folgt dass alle höheren Momente und Semiinvarianten existieren.

Mit Satz 2.2.8 können wir die Semiinvarianten aus den Momenten bzw. die Momente aus den Semiinvarianten berechnen. Aus statistischer Sicht lassen sich aus einer Stichprobe am besten die zentralen empirischen Momente schätzen. Beziehung (2.20) ermöglicht uns nun aus statistischen Schätzungen der zentrierten Momente eine Schätzfunktion für die Semiinvarianten abzuleiten.

Folgerung 2.2.9 Die k -te Semiinvariante der Zufallsgröße ξ lässt sich aus den ersten k zentrierten Momenten und $k - 1$ Semiinvarianten der Zufallsgröße ξ berechnen.

$$\begin{aligned}
s_1 &= m_1 = \mathbb{E}\xi \\
s_2 &= m_2 - m_1^2 = \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \mathbb{D}^2\xi \\
s_3 &= m_3 - 3m_2m_1 + 2m_1^3 = \mathbb{E}(\xi - \mathbb{E}\xi)^3 \\
s_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4 = \dots = \mathbb{E}(\xi - \mathbb{E}\xi)^4 - 3s_2^2 \\
s_5 &= m_5 - 5m_4m_1 - 10m_3m_2 + 20m_3m_1^2 + 30m_2^2m_1 - 60m_2m_1^3 + 24m_1^5 \\
&= \dots = \mathbb{E}(\xi - \mathbb{E}\xi)^5 - 10s_2s_3 \\
s_6 &= -120m_1^6 + 360m_2m_1^4 - 120m_3m_1^3 - 270m_1^2m_2^2 + 30m_1^2m_4 \\
&\quad + 120m_1m_2m_3 + 30m_2^3 - 15m_2m_4 - 10m_3^2 - 6m_5m_1 + m_6 \\
&= \dots = \mathbb{E}(\xi - \mathbb{E}\xi)^6 - 15s_2s_4 - 10s_3^2 - 15s_2^3 \\
s_7 &= 210m_1m_2m_4 - 1260m_1^2m_2m_3 + 210m_2^2m_3 - 21m_2m_5 - 35m_3m_4 \\
&\quad - 2520m_1^5m_2 + 840m_1^4m_3 + 2520m_1^3m_2^2 - 210m_1^3m_4 + 42m_1^2m_5 \\
&\quad - 630m_1m_2^3 + 140m_3^2m_1 - 7m_6m_1 + m_7 + 720m_1^7 \\
&= \dots = \mathbb{E}(\xi - \mathbb{E}\xi)^7 - 21s_2s_5 - 35s_3s_4 - 105s_3s_2^2 + 6s_1^6(s_1 - 1) \\
s_8 &= -5040m_1^8 + 20160m_1^6m_2 - 25200m_1^4m_2^2 + 1680m_4m_1^4 - 336m_1^3m_5 \\
&\quad + 10080m_1^2m_2^3 - 1680m_1^2m_3^2 + 56m_1^2m_6 - 630m_2^4 - 35m_4^2 - 8m_7m_1 \\
&\quad + m_8 + 13440m_3m_1^3m_2 - 2520m_4m_1^2m_2 - 5040m_1m_2^2m_3 + 420m_2^2m_4 \\
&\quad + 336m_2m_1m_5 + 560m_2m_3^2 - 28m_2m_6 + 560m_3m_1m_4 - 56m_3m_5 - 6720m_3m_1^5 \\
&= \dots = \mathbb{E}(\xi - \mathbb{E}\xi)^8 - 28s_2s_6 - 56s_3s_5 - 35s_4^2 - 210s_4s_2^2 - 280s_3^2s_2 \\
&\quad - 105s_2^4 \\
&\vdots \qquad \qquad \qquad \vdots
\end{aligned} \tag{2.21}$$

Insbesondere erhalten wir für den Erwartungswert $\mathbb{E}r_\tau = m_{\tau,1}(\underline{p}) = s_{\tau,1}(\underline{p}) = (\lambda\mu_Y + \tilde{\mu})\tau$ und die Varianz $\mathbb{D}^2r_\tau = s_{\tau,2}(\underline{p}) = ((\sigma_Y^2 + \mu_Y^2)\lambda + \sigma^2)\tau$. Durch Kenntnis der ersten Semiinvarianten können wir nun auch mit (2.21) die wichtigen statistischen Kennzahlen Schiefe und Exzess

$$\text{schiefe}(\underline{p}, \tau) = \frac{\mathbb{E}(r_\tau - \mathbb{E}r_\tau)^3}{\sqrt{(\mathbb{D}^2r_\tau)^3}} = \frac{s_{\tau,3}(\underline{p})}{\sqrt{s_{\tau,2}(\underline{p})^3}} = \frac{(3\sigma_Y^2 + \mu_Y^2)\lambda\mu_Y}{\sqrt{((\sigma_Y^2 + \mu_Y^2)\lambda + \sigma^2)^3\tau}} \tag{2.22}$$

$$\text{exzess}(\underline{p}, \tau) = \frac{\mathbb{E}(r_\tau - \mathbb{E}r_\tau)^4}{(\mathbb{D}^2r_\tau)^2} - 3 = \frac{s_{\tau,4}(\underline{p})}{s_{\tau,2}(\underline{p})^2} = \frac{(3\sigma_Y^4 + 6\sigma_Y^2\mu_Y^2 + \mu_Y^4)\lambda}{((\sigma_Y^2 + \mu_Y^2)\lambda + \sigma^2)^2\tau}, \tag{2.23}$$

deren Werte für normalverteilte Zufallsgrößen gerade null sind, explizit berechnen. Beide Kennzahlen sind dabei vom Lag τ abhängig und es gilt für alle festen Parameter $\underline{p} \in D$

$$\lim_{\tau \rightarrow 0+} \text{schiefe}(\underline{p}, \tau) = \infty \quad \text{bzw.} \quad \lim_{\tau \rightarrow 0+} \text{exzess}(\underline{p}, \tau) = \infty \quad (\lambda(\mu_Y^2 + \sigma_Y^2) \neq 0)$$

sowie

$$\lim_{\tau \rightarrow \infty} \text{schiefe}(\underline{p}, \tau) = 0 \quad \text{bzw.} \quad \lim_{\tau \rightarrow \infty} \text{exzess}(\underline{p}, \tau) = 0.$$

Dies deckt sich mit empirischen Beobachtungen, dass Returns mit einem geringen Lag deutlich von einer Normalverteilung abweichen, während für größere Lags die Abweichungen nicht mehr so stark auffallen.

Wegen Satz 2.2.8 enthalten die Momente und Semiinvarianten bis zur k -ten Ordnung die gleichen Informationen. Das folgende Lemma zeigt, dass wir von Semiinvarianten der Returns mit Lag τ_1 sofort auf Semiinvarianten der Returns mit Lag τ_2 schließen können.

Lemma 2.2.10 *Für alle Parametervektoren $\underline{p} \in D$ gilt für $\alpha > 0$*

$$s_{\alpha\tau,k}(\underline{p}) = \alpha s_{\tau,k}(\underline{p})$$

Beweis. Nach (2.18) gilt für alle $k \in \mathbb{N}$ und $\underline{p} \in D$ die Beziehung $s_{\tau,k}(\underline{p}) = \frac{\ln \varphi_{r_\tau}^{(k)}(0, \underline{p})}{i^k}$. Außerdem haben wir

$$\ln \varphi_{r_\tau}(\theta, \underline{p}) = \tau \ln \varphi_{r_1}(\theta, \underline{p}),$$

so dass für alle $\theta \in \mathbb{R}$

$$\frac{\partial^k \ln \varphi_{r_\tau}(\theta, \underline{p})}{\partial \theta^k} = \tau \frac{\partial^k \ln \varphi_{r_1}(\theta, \underline{p})}{\partial \theta^k}$$

und somit für $\alpha > 0$ wegen

$$\begin{aligned} s_{\alpha\tau,k}(\underline{p}) &= \frac{\ln \varphi_{r_{\alpha\tau}}^{(k)}(0, \underline{p})}{i^k} = \alpha\tau \frac{\ln \varphi_{r_1}^{(k)}(0, \underline{p})}{i^k} \\ &= \alpha \frac{\ln \varphi_{r_\tau}^{(k)}(0, \underline{p})}{i^k} = \alpha s_{\tau,k}(\underline{p}) \end{aligned}$$

die Behauptung folgt □

Lemma 2.2.10 ermöglicht uns nun beispielsweise von den Semiinvarianten der täglichen Returns mit $\alpha = 250$ zu den Semiinvarianten der jährlichen Returns überzugehen. Für die Momente gilt keine solche „Skalierungseigenschaft“, da beispielsweise aus der Beziehung $m_{\alpha\tau,2}(\underline{p}) = \alpha m_{\tau,2}(\underline{p})$ mit (2.21) sofort für alle $\alpha > 0$ die Beziehung

$$s_{\alpha\tau,2}(\underline{p}) = m_{\alpha\tau,2}(\underline{p}) - m_{\alpha\tau,1}(\underline{p})^2 = \alpha m_{\tau,2}(\underline{p}) - \alpha^2 m_{\tau,1}(\underline{p})^2 \neq \alpha s_{\tau,2}(\underline{p})$$

folgen würde, was offensichtlich im Widerspruch zu Lemma 2.2.10 steht. Wegen der „Skalierungseigenschaft“ werden wir in dieser Arbeit häufig die Semiinvarianten der Returns betrachten.

Bemerkung 2.2.11 Berechnet man für feste τ und $\underline{p} \in D$ die Momente nicht wie oben beschrieben, sondern direkt über die Definition

$$\begin{aligned} m_{\tau,k}(\underline{p}) &= \int_{-\infty}^{\infty} x^k f_{r_\tau}(x, \underline{p}) dx = \int_{-\infty}^{\infty} x^k \left(\sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) \right) dx \\ &= \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} \left(\int_{-\infty}^{\infty} x^k \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) dx \right), \end{aligned}$$

wird deutlich, dass das k -te Moment eine unendliche Überlagerung gewichteter k -ter Momente normalverteilter Zufallsgrößen $X_j \sim N(\tilde{\mu}\tau + j\mu_Y, \sigma^2\tau + j\sigma_Y^2)$, $j = 0, \dots, \infty$, ist. Die Gewichte $w_j = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!}$ stammen dabei aus der Poisson-Verteilung.

2.3 Spezialfall geometrische Brownsche Bewegung

Für den Fall

$$\lambda(\mu_Y^2 + \sigma_Y^2) = 0 \quad (2.24)$$

erhalten wir als Spezialfall der Jump-Diffusions-Prozesse das Kursmodell der geometrischen Brownschen Bewegung. Gleichung (2.24) ist für $\lambda = 0$ oder für $\mu_Y = 0$ sowie $\sigma_Y = 0$ erfüllt. Für $\lambda = 0$ treten wegen $\mathbb{P}(N_t = 0) = \exp(-\lambda t) = 1$ fast sicher keine Sprünge auf. Der zweite Fall, d.h. die mittlere Sprunghöhe μ_Y und die Standardabweichung der Sprunghöhen σ_Y verschwinden, impliziert $\ln Y_j \equiv 0$ (f.s.), so dass die Sprünge fast sicher keine Wirkung zeigen. Aus Bedingung (2.24) folgt wegen

$$\nu = \exp\left(\mu_Y + \frac{1}{2}\sigma_Y^2\right) - 1$$

unmittelbar $\lambda\nu = 0$. Die Trajektorien des Preisprozesses sind dann stetig, und die stochastische Differentialgleichung (2.1) vereinfacht sich zu

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

deren Lösung

$$S_t = S_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right)$$

wir mit Hilfe der Itô-Formel oder mit (2.24) in (2.7) erhalten. Die Returns r_τ entsprechen also der Brownschen Bewegung

$$r_\tau = \left(\mu - \frac{\sigma^2}{2}\right)\tau + \sigma W_\tau = \tilde{\mu}\tau + \sigma W_\tau.$$

Für festes τ ist der Return somit normalverteilt mit Erwartungswert $\tilde{\mu}\tau$ und Varianz $\sigma^2\tau$. Folglich entspricht die Dichtefunktion einfach nur der Dichte einer $N(\tilde{\mu}\tau, \sigma^2\tau)$ verteilten Zufallsgröße. Die Momente, Semiinvarianten und die charakteristische Funktion einer normalverteilten Zufallsgröße können ebenfalls leicht berechnet werden. Mit (2.24) erhalten wir

$$f_{r_\tau}(x, \underline{p}) = \frac{1}{\sqrt{\sigma^2\tau}} \phi\left(\frac{x - \tilde{\mu}\tau}{\sqrt{\sigma^2\tau}}\right) \quad \text{aus (2.15),}$$

$$\varphi_{r_\tau}(\theta, \underline{p}) = \exp\left(i\tilde{\mu}\tau\theta - \frac{\sigma^2}{2}\tau\theta^2\right) \quad \text{aus (2.16),}$$

$$s_{\tau,1}(\underline{p}) = \tilde{\mu}\tau, \quad s_{\tau,2}(\underline{p}) = \sigma^2\tau \quad \text{und} \quad s_{\tau,k}(\underline{p}) = 0 \quad (k \geq 3) \quad \text{aus (2.19) sowie}$$

$$\text{schiefe}(\underline{p}, \tau) = 0 \quad \text{und} \quad \text{exzess}(\underline{p}, \tau) = 0 \quad \text{aus (2.22) und (2.23).}$$

Für diesen Spezialfall verschwindet also die unendliche Überlagerung normalverteilter Zufallsgrößen, da der Return selber normalverteilt ist. Insbesondere ist dann das Problem der Parameterschätzung korrekt (gut gestellt), wenn a priori bekannt ist, dass die Aktienkurse durch eine geometrische Brownsche Bewegung beschrieben werden.

Kapitel 3

Parameterschätzung – das inverse Problem

3.1 Übergang von Kursdaten zu Returndaten

In diesem Kapitel diskutieren wir verschiedene Zugänge zur Schätzung der Parameter eines Jump-Diffusions-Prozesses. Wir werden später in Kapitel 5 zeigen, dass das inverse Problem der Parameterbestimmung für Jump-Diffusions-Prozesse inkorrekt ist und diese Inkorrektheit unabhängig vom gewählten Zugang zu Instabilitätsproblemen bei der Parameteridentifizierung führt. Die Inkorrektheit wird dadurch verursacht, dass verschiedene Parameterkonstellationen ein ähnliches Kursverhalten beschreiben können.

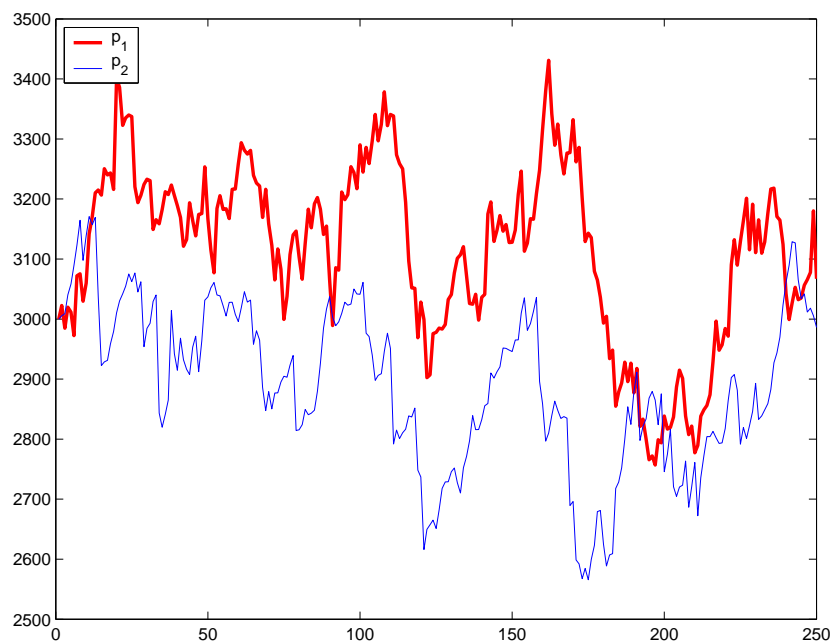


Abbildung 3.1: Trajektorien eines Jump-Diffusions-Prozesses mit verschiedenen Parameterkonstellationen

Ein Beispiel dafür ist in Abbildung 3.1 zu sehen, in der eine Trajektorie $S^*(\underline{p}_1)$ für den Parametervektor $\underline{p}_1 = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y) = (0.15, 0.25, 3, 0.05, 0.05)$ und eine Trajektorie $S^*(\underline{p}_2)$ für den Parametervektor $\underline{p}_2 = (0.1, 0.1, 100, -0.01125, 0.02)$ abgebildet ist. Der Parametervektor \underline{p}_1 weist mit $\lambda = 3$ eine geringe Sprunghäufigkeit und mit $\sigma = 0.25$ eine hohe Volatilität des Diffusionsteils auf. Hingegen ist die Volatilität des Diffusionsterms des zweiten Parametervektors \underline{p}_2 mit $\sigma = 0.1$ eher gering. Wegen der hohen Sprunghäufigkeit mit $\lambda = 100$ steigt aber im Falle von \underline{p}_2 die Gesamtvolatilität des Kurses, wobei die Kurssprünge wegen $\mu_Y = -0.01125$ und $\sigma_Y = 0.02$ nur gering ausfallen¹. Allgemein bereitet es Schwierigkeiten, viele kleine Sprünge von einer hohen Volatilität des Diffusionsteils in einer Trajektorie zu unterscheiden. In Kapitel 2 haben wir bereits motiviert, dass es sinnvoll ist die Verteilung des stochastischen Prozesses zu berücksichtigen. Ferner haben wir in Abschnitt 2.2 dargestellt, dass es für Jump-Diffusions-Prozesse vorteilhaft ist, von den äquidistanten Kursdaten $S^* = (S_{t_0}^*, \dots, S_{t_n}^*)$ zu den stationären Returns $r^* = (r_1^*, \dots, r_n^*)$ (siehe Formel (2.12)) überzugehen, da diese gerade n unabhängige Realisierungen der Zufallsgröße r_τ liefern. Abbildung 3.2 zeigt die zu den Trajektorien $S^*(\underline{p}_1)$ und $S^*(\underline{p}_2)$ gehörenden Returns $r^*(\underline{p}_1)$ bzw. $r^*(\underline{p}_2)$. Auch hier können wir kein grundsätzlich verschiedenes Verhalten der Trajektorien beobachten, obwohl geringe Unterschiede im Vergleich zu den Preistrajektorien erkennbar werden.

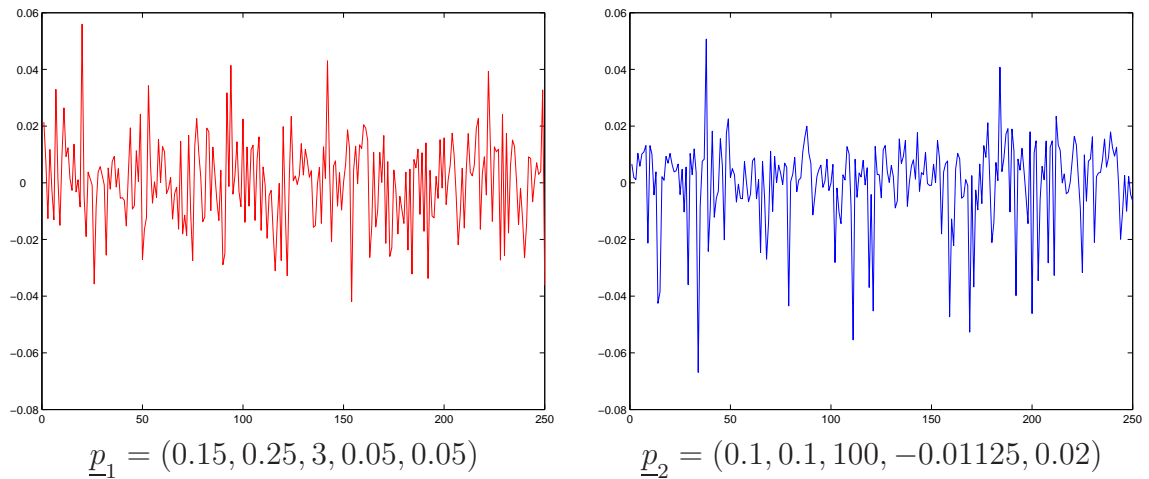
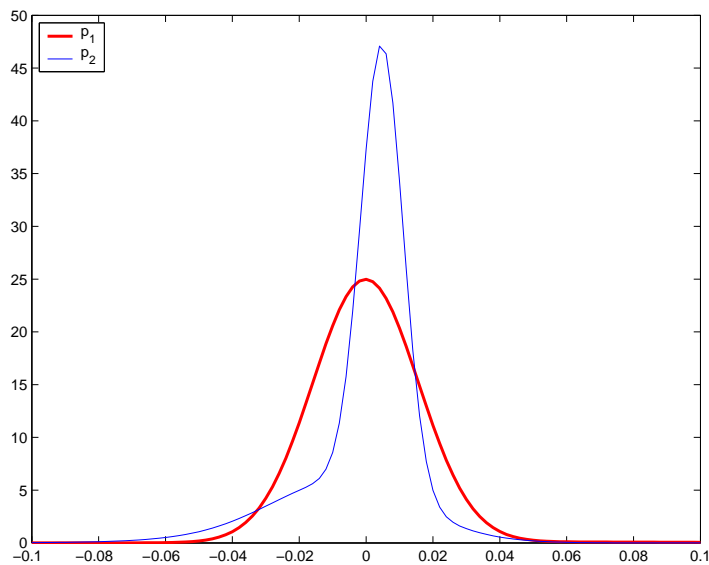


Abbildung 3.2: Returns eines Jump-Diffusions-Prozesses

Wir kennen von der absolut stetigen Zufallsgröße r_τ die Verteilung, insbesondere die Dichtefunktion (2.15) sowie die Semiinvarianten (2.19).

Abbildung 3.3 zeigt die Dichtefunktionen der täglichen Returns ($\tau = \frac{1}{250}$) für die Parametervektoren \underline{p}_1 bzw. \underline{p}_2 . Obwohl wir Schwierigkeiten haben den Trajektorien aus Abbildung 3.1 die Parametervektoren \underline{p}_1 und \underline{p}_2 zu zuordnen, unterscheiden sich die Dichtefunktionen $f_{r_\tau}(\cdot, \underline{p}_1)$ und $f_{r_\tau}(\cdot, \underline{p}_2)$ recht deutlich. Zur Parameterbestimmung steht uns aber üblicherweise nur eine Trajektorie zur Verfügung. Das betrachtete Beispiel hat gezeigt, dass verschiedene Parametervektoren, deren Dichtefunktionen sich deutlich

¹Für die mittlere Sprunghöhe $\nu = \mathbb{E}(Y_j - 1)$ (vgl. Kapitel 2) gilt gerade $\nu = \exp(\mu_Y + \frac{1}{2}\sigma_Y^2) - 1 = -0.01$, d.h. die Kurse fallen dann wegen $S_{T_j^+} = S_{T_j^-} + (Y_j - 1)S_{T_j^-} = Y_j S_{T_j^-}$ im Mittel um ein Prozent.

Abbildung 3.3: Dichten der Returns für \underline{p}_1 und \underline{p}_2

unterscheiden, recht ähnliche Trajektorien erzeugen können. Daher ist zu befürchten, dass die aus $r^*(\underline{p}_1)$ bzw. $r^*(\underline{p}_2)$ berechneten empirischen Dichtefunktionen sich ähneln können.

Im Weiteren werden wir die folgende Definition der empirischen Dichtefunktion in Anlehnung an [62] verwenden, um aus r^* die empirische Dichtefunktion zu berechnen.

Definition 3.1.1 *Es seien I_1, I_2, \dots gleichlange Intervalle der Länge $l = |I_q|$. Wir betrachten die Zerlegung $\mathcal{I} = (I_q)_{q \in \mathbb{N}}$ mit $\mathbb{R} = \bigcup_{q \in \mathbb{N}} I_q$ und $I_q \cap I_p = \emptyset$. Für die Stichprobe (x_1, x_2, \dots, x_n) definieren wir die empirische Dichtefunktion*

$$h = h(\mathcal{I}; x_1, x_2, \dots, x_n) = \sum_q \frac{\#\{i : x_i \in I_q\}}{n|I_q|} \mathbb{I}_{I_q}. \quad (3.1)$$

Da wir hier nur äquidistante Zerlegungen \mathcal{I} betrachten, stimmt der Durchmesser der Zerlegung $|\mathcal{I}|$ grundsätzlich mit $l = |I_q|$ überein. Der folgende Satz aus [62, Theorem VII.13.2.] sichert uns nun die Konvergenz (in Wahrscheinlichkeit) der empirischen Dichtefunktion gegen die Dichtefunktion, falls wir die Trajektorie S^* genügend lange beobachten², d.h. beliebig viele Returns $r^* = (r_1^*, r_2^*, \dots)$ mit Lag τ zur Bestimmung der empirischen Dichte h nutzen.

Satz 3.1.2 *Es sei X_1, X_2, \dots eine Folge unabhängiger Zufallsgrößen, die identisch wie die absolut stetigen Zufallsgröße X mit Dichtefunktion f verteilt sind. Ferner sei $\mathcal{I}_1, \mathcal{I}_2, \dots$ eine Folge von Zerlegungen von \mathbb{R} , so dass*

$$|\mathcal{I}_n| = |I_q(n)| \longrightarrow 0 \quad \text{und} \quad n |\mathcal{I}_n| \longrightarrow +\infty \quad \text{für } n \rightarrow \infty.$$

²Selbst wenn wir alle Werte der Trajektorie bis zu einem Zeitpunkt $\bar{T} < \infty$ kennen würden und demnach überabzählbar viele Kursdaten zur Verfügung hätten, könnten wir nur für $\tau \rightarrow 0$ beliebig viele Returns (mit Lag τ) aus den Kursdaten berechnen.

Dann gilt für jeden Stetigkeitspunkt x von f

$$\lim_{n \rightarrow \infty} h(\mathcal{I}_n; X_1, \dots, X_n)(x) = f(x) \quad \text{in Wahrscheinlichkeit.}$$

Da für fixiertes $\underline{p} \in D$ die Dichtefunktion

$$f_{r_\tau}(x, \underline{p}) = \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right)$$

für alle $x \in \mathbb{R}$ stetig ist, folgt die Konvergenz (in Wahrscheinlichkeit) der empirischen Dichtefunktion $h = h(r^*)$ gegen f_{r_τ} für alle $x \in \mathbb{R}$. Somit können wir die empirische Dichtefunktion $h(r^*)$ als eine Störung der Funktion f_{r_τ} interpretieren.

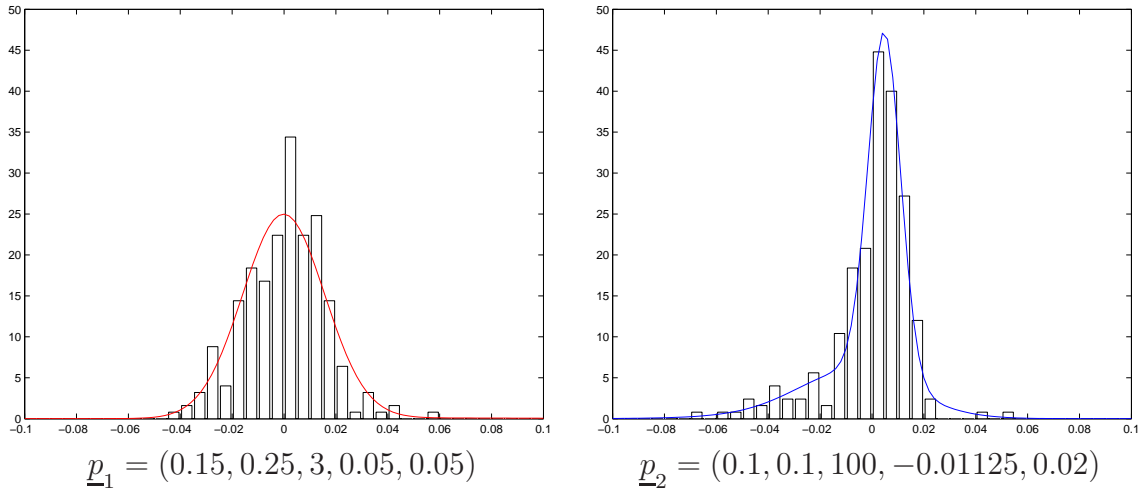


Abbildung 3.4: Dichtefunktionen und Histogramme

In Abbildung 3.4 sind für unser Beispiel die aus $r^*(\underline{p}_1)$ bzw. $r^*(\underline{p}_2)$ bestimmten empirischen Dichtefunktionen und die entsprechenden Dichtefunktionen der Parametervektoren \underline{p}_1 und \underline{p}_2 dargestellt. Dabei zeigt sich nun der zufallsbedingte glättende Charakter des Vorwärtsproblems. Die empirischen Dichtefunktionen (siehe linke Graphik von Abbildung 3.5) unterscheiden sich in diesem Beispiel nicht mehr so deutlich. Natürlich treten diese Abweichungen, wie die rechte Graphik von Abbildung 3.5 mit den empirischen Dichtefunktionen einer weiteren Simulation zeigt, zufällig auf. Bei dieser Simulation, mit den gleichen Parametervektoren, können wir auch die empirischen Dichtefunktionen deutlich unterscheiden. Wir könnten aber wiederum andere Parametervektoren finden, so dass deren simulierten Trajektorien ähnliche empirische Dichtefunktionen wie in der rechten Graphik von Abbildung 3.5 implizieren. Wir werden an dieser Stelle nicht nach solchen Vektoren suchen, da dieses Beispiel lediglich den zufallsbedingten glättenden Charakter des Vorwärtsproblems illustrieren sollte.

Bemerkung 3.1.3 Wir werden später in Kapitel 4 und Kapitel 5 zeigen, dass nicht nur die oben illustrierte zufallsbedingt glättende Eigenschaft des Vorwärtsproblems die Inkorrektheit des inversen Problems verursacht. So können wir Parameterkonstellationen

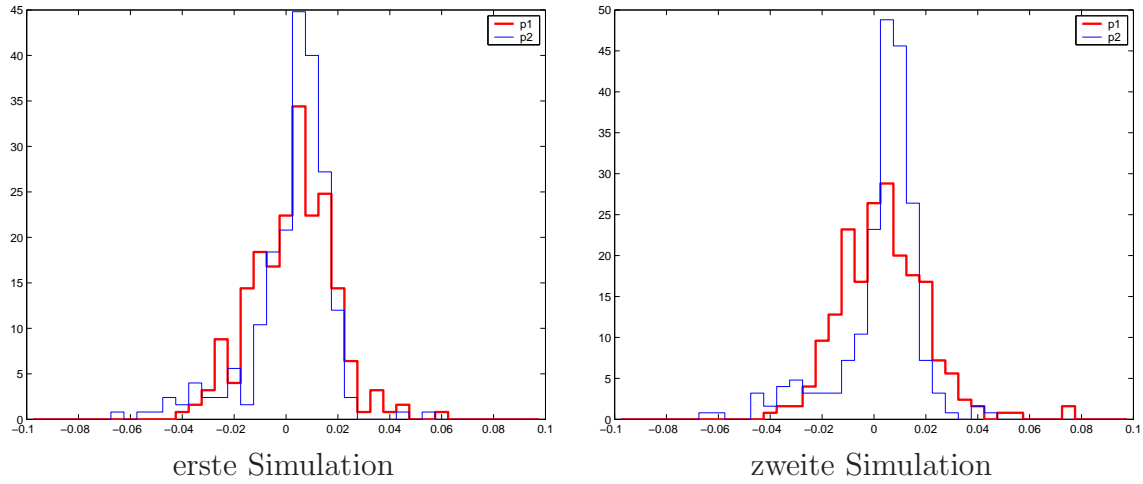


Abbildung 3.5: Empirische Dichten aus $r^*(\underline{p}_1)$ und $r^*(\underline{p}_2)$ für verschiedene Simulationen

finden, deren Abstand beliebig groß wird, obwohl sich die entsprechenden Dichtefunktionen kaum unterscheiden. Folglich wird sich dann auch ein Großteil der Trajektorien, die von den jeweiligen Parametervektoren beeinflusst werden, nicht wesentlich unterscheiden.

Wir werden nun im Weiteren nur noch die Returns r^* betrachten, da nach Bemerkung 2.2.3 eine eindeutige Beziehung zwischen Returns und Aktienkursen besteht. Dabei ist r^* eine Realisierung des Zufallsvektors (2.13). Nutzen wir zur Bestimmung der Parameter empirische Daten $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ (vgl. (2.9)) anstatt exakter Daten r^* , so betrachten wir für einen geeignet gewählten Abstandsbegriff $d(\cdot, \cdot)$ zwischen dem vom Parameter \underline{p} abhängigen Vektor r^* der Returns und dem Vektor r^δ der empirischen Returns das Extremalproblem

$$d(r^\delta, r^*(\underline{p})) \longrightarrow \min \quad \text{für } \underline{p} \in D. \quad (3.2)$$

Bemerkung 3.1.4 Da im Falle von empirischen Daten die konkrete Trajektorie S^* bzw. das zugehörige $\omega \in \Omega$ und somit auch die exakten Returns r^* nicht bekannt sind, können wir zur Bildung des Abstandsbegriffs $d(r^\delta, r^*)$ beispielsweise den Zufallsvektor $\underline{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ verwenden, in welchem die Zufallsgrößen ξ_i unabhängig und identisch wie r_τ verteilt sind. Der Vektor r^* ist dann gerade eine Realisierung des Zufallsvektors $\underline{\xi}$, der genau wie r^* vom Parametervektor $\underline{p} \in D$ beeinflusst wird. Bei der Verwendung von $\underline{\xi}$ anstelle von r^* zur Bildung des Abstandsbegriffs $d(r^\delta, r^*)$ schreiben wir zur besseren Unterscheidung $\tilde{d}(r^\delta, r^*)$. Als möglicher Abstandsbegriff wäre dann beispielsweise

$$\tilde{d}(r^\delta, r^*) := d(r^\delta, \mathbb{E}\underline{\xi}) = \|r^\delta - \mathbb{E}\underline{\xi}\|^2 \quad (3.3)$$

oder

$$\tilde{d}(r^\delta, r^*) := \mathbb{E} d(r^\delta, \underline{\xi}) = \mathbb{E} \|r^\delta - \underline{\xi}\|^2 \quad (3.4)$$

denkbar, wobei wir die euklidische Norm verwenden³. Die Abstände (3.3) und (3.4)

³Falls nicht anders angegeben, werden wir in dieser Arbeit den Abstand zwischen zwei Vektoren aus dem \mathbb{R}^n immer mit der euklidischen Norm $\|\cdot\|$ messen.

lassen sich mit

$$\begin{aligned}\mathbb{E} \|r^\delta - \underline{\xi}\|^2 &= \mathbb{E} \left(\sum_{i=1}^n (\xi_i - r_i^\delta)^2 \right) = \sum_{i=1}^n \mathbb{E} (\xi_i - r_i^\delta)^2 \\ &= \sum_{i=1}^n ((\mathbb{E}\xi_i)^2 - 2r_i^\delta \mathbb{E}\xi_i + (r_i^\delta)^2 + \mathbb{E}\xi_i^2 - (\mathbb{E}\xi_i)^2) = \|r^\delta - \mathbb{E}\underline{\xi}\|^2 + n\mathbb{D}^2 r_\tau\end{aligned}$$

ineinander umrechnen.

Bemerkung 3.1.5 Zur Lösung des inversen Problems der Parameteridentifikation könnten wir wegen Bemerkung 3.1.4 und $\mathbb{E}\xi_i = \mathbb{E}r_\tau = m_{\tau,1}(\underline{p})$ und $\mathbb{D}^2 r_\tau = s_{\tau,2}(\underline{p})$ die Least-Squares Probleme

$$F_1(\underline{p}, r^\delta) = \|r^\delta - \mathbb{E}\underline{\xi}\|^2 = \sum_{i=1}^n (r_i^\delta - m_{\tau,1}(\underline{p}))^2 \longrightarrow \min \quad \text{für } \underline{p} \in D \quad (3.5)$$

bzw.

$$F_2(\underline{p}, r^\delta) = \mathbb{E} \|r^\delta - \underline{\xi}\|^2 = F_1(\underline{p}, r^\delta) + n s_{\tau,2}(\underline{p}) \longrightarrow \min \quad \text{für } \underline{p} \in D \quad (3.6)$$

betrachten. Allerdings sind die Probleme nicht geeignet, da obige Formulierungen nur ungenügend Informationen zur Parameterbestimmung enthalten. Für innere Punkte $\underline{p} \in D$ von Problem (3.5) erhalten wir aus

$$\frac{\partial F_1(\underline{p}, r^\delta)}{\partial p_i} = 0 \quad (i = 1, 2, \dots, 5)$$

nur die notwendige Bedingung $m_{\tau,1}(\underline{p}) = \frac{1}{n} \sum_{i=1}^n r_i^\delta$, die nicht ausreicht den Parametervektor $\underline{p} \in D$ zu bestimmen. Wir können die Funktion $F_1(\underline{p}, r^\delta)$ mit $x := m_{\tau,1}(\underline{p})$, $a := \sum_{i=1}^n r_i^\delta$ und $b := \sum_{i=1}^n (r_i^\delta)^2 > 0$ auch wie folgt umformulieren

$$\begin{aligned}F_1(\underline{p}, r^\delta) &= \sum_{i=1}^n (r_i^\delta - m_{\tau,1}(\underline{p}))^2 = \sum_{i=1}^n (r_i^\delta)^2 - 2m_{\tau,1}(\underline{p}) \sum_{i=1}^n r_i^\delta + n m_{\tau,1}(\underline{p})^2 \\ &= nx^2 - 2ax + b.\end{aligned}$$

Da das Minimum der quadratischen Funktion $nx^2 - 2ax + b$ nicht für $x = \pm\infty$ angenommen wird, können wir auch ausschließen, dass Randpunkte $\underline{p} \in D$ mit $m_{\tau,1}(\underline{p}) = \pm\infty$ zur Lösung von (3.5) gehören. Die Parametervektoren mit $\sigma = 0$, $\lambda = 0$ oder $\sigma_Y = 0$ liefern auch keine zusätzlichen Bedingungen, so dass wir \underline{p}^δ nicht bestimmen können. Das Optimierungsproblem (3.6) hat im Inneren von D überhaupt keine Lösung, da

$$\frac{\partial F_2(\underline{p}, r^\delta)}{\partial p_i} = 0 \quad (i = 1, 2, \dots, 5)$$

zu widersprüchlichen Bedingungen führt. Für Parametervektoren $\underline{p} \in D$ mit $\mathbb{D}^2 r_\tau = s_{\tau,2}(\underline{p}) \rightarrow 0$ strebt $F_2(\underline{p}, r^\delta)$ gegen das Infimum. Der Fall $s_{\tau,2}(\underline{p}) = ((\sigma_Y^2 + \mu_Y^2)\lambda + \sigma^2)\tau \rightarrow 0$ tritt aber gerade für $\sigma \rightarrow 0$ und $(\sigma_Y^2 + \mu_Y^2)\lambda \rightarrow 0$ unabhängig von r^δ ein, so dass es nicht sinnvoll ist, \underline{p}^δ als Lösung von (3.5) oder (3.6) zu bestimmen.

Da für Returns mit Lag τ sowohl die Verteilungsdichte als auch die Momente und Semiinvarianten bekannt sind, bieten sich auch herkömmliche statistische Methoden⁴ zur Parameterschätzung aus den Daten $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ an. Obwohl die Disziplinen Mathematische Statistik und Theorie der Inversen Probleme in ihrer Methodik teilweise recht verschieden sind, gibt es trotz der abweichenden „mathematischen Sprechweisen“ viele Gemeinsamkeiten (siehe z.B. [26], [36, Ch. 5], [38, Abschnitt 4.1.6]). Natürlich unterscheiden sich aber die Herangehensweisen an die Lösung des inversen Problems der Parameteridentifikation und teilweise auch die Zielstellungen. Der Statistiker ist an konsistenten, erwartungstreuen und wirksamen Schätzfunktionen für die gesuchten Parameter sowie deren Verteilung interessiert. Bei der inversen Theorie analysiert man zunächst die Inkorrektheit des Problems und sucht anschließend nach stabilen näherungsweise Lösungen, um die Wirkung der Inkorrektheit weitgehend einzudämmen. Datenfehler sollten trotz Inkorrektheit keinen großen Einfluss auf die Lösung haben. Dies erreicht man, indem das inkorrekte inverse Problem regularisiert wird und anstatt der inkorrekten Aufgabe eine stabile Ersatzaufgabe gelöst wird. Die Regularisierung wird dabei in der Regel durch einen skalaren Regularisierungsparameter $\alpha > 0$ gesteuert. Von einer „guten“ Regularisierung erwartet man, dass die regularisierte Lösung gegen die exakte Lösung konvergiert, wenn der Datenfehler δ gegen null konvergiert. Für unser konkretes Problem der Parameterschätzung entspricht die Tendenz $\delta \rightarrow 0$ wegen des Gesetzes der großen Zahlen gerade der Situation, dass der Stichprobenumfang n der Stichprobe $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ über alle Grenzen wächst. In dieser Betrachtungsweise stimmen die Begriffe Stabilität und Konsistenz im Prinzip überein (siehe [26]). Da man in praxi zur Schätzung der Parameter nicht beliebig viele Daten zur Verfügung hat, scheint das Ziel der inversen Theorie, eine bezüglich der Daten $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ stabile Lösung zu berechnen, eine gangbare Alternative zu sein.

Wir werden in Abschnitt 3.2 zunächst kurz auf die statistischen Standardzugänge eingehen und anschließend einen analytischen Zugang unter Einbeziehung von Methoden der inversen Theorie zur Lösung unseres allgemeinen inversen Problems diskutieren.

3.2 Statistischer Zugang zur Parameterschätzung

Es gibt viele Arbeiten zur Parameterschätzung (z.B. [3], [41], [46], [48]), die Modifikationen statistischer Verfahren zur Parameterschätzung in Jump-Diffusions-Prozessen oder anderer stochastischer Prozesse mit einem Sprungteil vorstellen. Andere Verfahren, auf die wir im Weiteren nicht eingehen werden, sind die Methode der simulierten Momente (siehe z.B. [16], [30]) oder die Methode der Markov-Chain-Monte-Carlo-Simulation (MCMC) (siehe z.B. [24], [25]). In [12] findet man einen allgemeinen Überblick zur Parameterschätzung für verschiedene Sprung-Prozesse.

⁴Man kann auch die statistischen Methoden als eine Lösung des Least-Squares Problems (3.2) mit entsprechend gewählten Abstandsbegriff $d(r^\delta, r^*(p))$ auffassen. So kann beispielsweise für die Momentenmethode $d(r^\delta, r^*(p))$ als euklidische Norm der Differenz der ersten k Momente $(m_{\tau,1}(p), \dots, m_{\tau,k}(p))$ und der ersten k von r^δ abgeleiteten empirischen Momente definiert werden.

3.2.1 Maximum-Likelihood-Methode

Da für jeden Parametervektor $\underline{p} \in D$ die Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ (vgl. (2.15)) für die Returns mit Lag τ bekannt ist, können wir mit Hilfe der Maximum-Likelihood-Methode den Parametervektor $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in D$ des Jump-Diffusions-Prozesses aus der Stichprobe $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ schätzen. Wir wählen als Schätzwert dazu den Parametervektor $\underline{p}^\delta \in D$, der die Likelihood-Funktion der Stichprobe r^δ maximiert und betrachten das Optimierungsproblem

$$L(r^\delta | \underline{p}) = \prod_{i=1}^n f_{r_\tau}(r_i^\delta, \underline{p}) \longrightarrow \max \quad \text{für } \underline{p} \in D. \quad (3.7)$$

In vielen praktischen Fällen geht man zur logarithmierten Likelihood-Funktion über und löst die Extremwertaufgabe

$$\ln L(r^\delta | \underline{p}) = \sum_{i=1}^n \ln f_{r_\tau}(r_i^\delta, \underline{p}) \longrightarrow \max \quad \text{für } \underline{p} \in D \quad (3.8)$$

anstelle von (3.7), da die Optima beider Funktionen übereinstimmen.

Lemma 3.2.1 *Die logarithmierte Likelihood-Funktion der Stichprobe r^δ für den Return r_τ mit Lag τ eines Jump-Diffusions-Prozesses hat die folgende Gestalt*

$$\ln L(r^\delta | \underline{p}) = -n\lambda\tau - \frac{n}{2} \ln \frac{\pi}{2} + \sum_{i=1}^n \ln \left(\sum_{j=0}^{\infty} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp \left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)} \right) \right). \quad (3.9)$$

Für innere Punkte \underline{p} des Definitionsbereichs D können wir den Gradienten der logarithmierten Likelihood-Funktion betrachten. Die Randpunkte müssen dann noch gesondert betrachtet werden. Zur Lösung von (3.8) erhalten wir als notwendige Bedingungen des Extremums

$$\frac{\partial \ln L(r^\delta | \underline{p})}{\partial p_i} = \sum_{i=1}^n \frac{1}{f_{r_\tau}(r_i^\delta, \underline{p})} \frac{\partial f_{r_\tau}(r_i^\delta, \underline{p})}{\partial p_i} = 0 \quad (i = 1, \dots, 5)$$

und damit ein homogenes, nichtlineares Gleichungssystem mit folgender linker Seite

$$\begin{aligned} \frac{\partial \ln L(r^\delta | \underline{p})}{\partial \mu} &= \sum_{i=1}^n \frac{\sum_{j=0}^{\infty} \frac{(\lambda\tau)^j \exp \left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)} \right)}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \frac{\tau(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)}{(\sigma^2\tau + j\sigma_Y^2)}}{\sum_{j=0}^{\infty} \frac{(\lambda\tau)^j}{j!} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp \left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)} \right)} \\ \frac{\partial \ln L(r^\delta | \underline{p})}{\partial \sigma^2} &= \sum_{i=1}^n \frac{\sum_{j=0}^{\infty} \frac{(\lambda\tau)^j \exp \left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)} \right)}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \frac{-\tau \left(1 + (r_i^\delta - \tilde{\mu}\tau - j\mu_Y) \left(1 - \frac{r_i^\delta - \tilde{\mu}\tau - j\mu_Y}{\sigma^2\tau + j\sigma_Y^2} \right) \right)}{2(\sigma^2\tau + j\sigma_Y^2)}}{\sum_{j=0}^{\infty} \frac{(\lambda\tau)^j}{j!} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp \left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)} \right)} \end{aligned}$$

$$\begin{aligned}
\frac{\partial \ln L(r^\delta | \underline{p})}{\partial \lambda} &= \sum_{i=1}^n \frac{\sum_{j=0}^{\infty} \frac{(\lambda \tau)^j \exp\left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right)}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \tau \left(\frac{j}{\lambda\tau} - 1 - \nu \frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)}{(\sigma^2\tau + j\sigma_Y^2)}\right)}{\sum_{j=0}^{\infty} \frac{(\lambda \tau)^j}{j!} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp\left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right)} \quad (3.10) \\
\frac{\partial \ln L(r^\delta | \underline{p})}{\partial \mu_Y} &= \sum_{i=1}^n \frac{\sum_{j=0}^{\infty} \frac{(\lambda \tau)^j \exp\left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right)}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)}{(\sigma^2\tau + j\sigma_Y^2)} (j - \lambda\tau(\nu + 1))}{\sum_{j=0}^{\infty} \frac{(\lambda \tau)^j}{j!} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp\left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right)} \\
\frac{\partial \ln L(r^\delta | \underline{p})}{\partial \sigma_Y^2} &= \sum_{i=1}^n \frac{\sum_{j=0}^{\infty} \frac{(\lambda \tau)^j \exp\left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right)}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \frac{-\left(j + (r_i^\delta - \tilde{\mu}\tau - j\mu_Y)(\lambda\tau(\nu + 1) - j \frac{r_i^\delta - \tilde{\mu}\tau - j\mu_Y}{\sigma^2\tau + j\sigma_Y^2})\right)}{2(\sigma^2\tau + j\sigma_Y^2)}}{\sum_{j=0}^{\infty} \frac{(\lambda \tau)^j}{j!} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp\left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right)}.
\end{aligned}$$

Für das nichtlineare Gleichungssystem (3.10) lässt sich wohl keine explizite Lösung angeben. Allerdings kann man versuchen, dieses Problem numerisch zu lösen, was aber aufgrund der komplexen Struktur problematisch ist. Da die Dichtefunktion (2.15)

$$f_{r_\tau}(x, \underline{p}) = \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right)$$

eine unendliche Summe gewichteter Dichtefunktionen normalverteilter Zufallsgrößen mit Erwartungswert $\tilde{\mu}\tau + j\mu_Y$ und Varianz $\sigma^2\tau + j\sigma_Y^2$ ist, können wir zur numerischen Umsetzung nur endlich viele Summanden berücksichtigen. In der Literatur ([55], [41], [12]) nutzt man häufig die folgende Approximation der Dichtefunktion

$$f(x, \underline{p}) \approx (1 - \lambda\tau) \frac{1}{\sqrt{\sigma^2\tau}} \phi\left(\frac{x - \tilde{\mu}\tau}{\sqrt{\sigma^2\tau}}\right) + \lambda\tau \frac{1}{\sqrt{\sigma^2\tau + \sigma_Y^2}} \phi\left(\frac{x - \tilde{\mu}\tau - \mu_Y}{\sqrt{\sigma^2\tau + \sigma_Y^2}}\right), \quad (3.11)$$

um die Maximierung der Likelihood-Funktion (3.7) und (3.8) bzw. das davon abgeleitete nichtlineare Gleichungssystem (3.10) einfacher zu lösen. Man erhält (3.11), indem man die Gewichte $w_j = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!}$ der Dichtefunktion (2.15) in eine Taylorreihe zweiter Ordnung im Nullpunkt entwickelt⁵. Daher liefert (3.11) nur für den Fall $\lambda\tau \ll 1$ gute Resultate. Im Wesentlichen entspricht die Approximation (3.11) gerade einem Abschneiden der unendlichen Summe der Dichtefunktion (2.15) nach den ersten zwei Summanden. Eine im Allgemeinen bessere Approximation erhalten wir, indem wir die ersten $N+1$ Summanden der Dichtefunktion berücksichtigen. Wir betrachten dann also anstelle von $f_{r_\tau}(\cdot, \underline{p})$ die Approximation

$$f_{r_\tau}^N(x, \underline{p}) := \sum_{j=0}^N e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right). \quad (3.12)$$

⁵Es gilt $w_0 = e^{-\lambda\tau} = 1 - \lambda\tau + \mathcal{O}((\lambda\tau)^2)$, $w_1 = e^{-\lambda\tau} \lambda\tau = \lambda\tau + \mathcal{O}((\lambda\tau)^2)$ und für $j \geq 2$ gerade $w_j = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} = \frac{\mathcal{O}((\lambda\tau)^2)}{j!}$.

Wegen der Nichtnegativität aller Reihenglieder der Dichtefunktion erhalten wir für den Approximationsfehler

$$\begin{aligned} R^N(x, \underline{p}) &= f_{r_\tau}(\cdot, \underline{p}) - f_{r_\tau}^N(\cdot, \underline{p}) \\ &= \sum_{j=N+1}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right). \end{aligned}$$

Lemma 3.2.2 Für den Approximationsfehler R^N gilt für alle $x \in \mathbb{R}$ und $\underline{p} \in D$ die folgende Abschätzung

$$R^N(x, \underline{p}) = |f_{r_\tau}(x, \underline{p}) - f_{r_\tau}^N(x, \underline{p})| \leq \frac{1}{\sqrt{2\pi\sigma^2\tau}} \left(1 - e^{-\lambda\tau} \sum_{j=0}^N \frac{(\lambda\tau)^j}{j!}\right). \quad (3.13)$$

Beweis. Wir erhalten die Ungleichung (3.13) wegen $\frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \leq \frac{1}{\sqrt{\sigma^2\tau}}$ sowie $\phi(x) \leq \frac{1}{\sqrt{2\pi}}$ durch folgende Abschätzungen

$$\begin{aligned} R^N(x, \underline{p}) &= \sum_{j=N+1}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) \\ &\leq \frac{1}{\sqrt{2\pi\sigma^2\tau}} \left(e^{-\lambda\tau} \sum_{j=0}^{\infty} \frac{(\lambda\tau)^j}{j!} - e^{-\lambda\tau} \sum_{j=0}^N \frac{(\lambda\tau)^j}{j!}\right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2\tau}} \left(1 - e^{-\lambda\tau} \sum_{j=0}^N \frac{(\lambda\tau)^j}{j!}\right) \end{aligned}$$

□

Da für alle Parameter $\underline{p} \in D$ der Fehler R^N bezüglich N monoton fallend ist und für $N \rightarrow \infty$ gegen null konvergiert, brauchen wir für den Abbruch nur ein genügend großes $N = N(\underline{p}, \varepsilon)$ für $R^N(x, \underline{p}) < \varepsilon$ zu wählen. In vielen praktischen Fällen erhält man für $N = 100$ akzeptable Fehler beim Abschneiden (siehe auch [41]).

Für die logarithmierte Likelihood-Funktion der abgeschnittenen Dichtefunktion gilt analog zu (3.9) die Beziehung

$$\begin{aligned} \ln L^N(r^\delta | \underline{p}) &= \sum_{i=1}^n \ln f^N(r_i^\delta, \underline{p}) = -n\lambda\tau - \frac{n}{2} \ln \frac{\pi}{2} \\ &\quad + \sum_{i=1}^n \ln \left(\sum_{j=0}^N \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp\left(-\frac{(r_i^\delta - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right) \right). \end{aligned} \quad (3.14)$$

Analog erhalten wir auch obiges Gleichungssystem des Gradienten der logarithmierten Likelihood-Funktion von $f^N(\cdot, \underline{p})$, indem wir die unendlichen Summen in (3.10) nach $N+1$ Summationen abschneiden. Versucht man nun dieses Gleichungssystem numerisch zu lösen, so ergeben sich aufgrund der Struktur von (3.10) eine Vielzahl numerischer

Schwierigkeiten. Die Arbeit [41] „Pitfalls in Estimating Jump-Diffusion Models“ von Honoré zeigt ebenfalls, dass bei der Schätzung der Parameter eines Jump-Diffusions-Prozesses mittels der herkömmlichen Maximum-Likelihood-Methode schwerwiegende Probleme auftreten können, da für einige Parameter die Likelihood-Funktion unbeschränkt ist. Eine besondere Rolle spielt dabei der Fall $\sigma \rightarrow 0$, den wir in Abschnitt 4.3 noch genauer untersuchen werden. Honoré schlägt eine Einschränkung des in Abschnitt 2.1 eingeführten Definitionsbereichs $D = \{\underline{p} \in \mathbb{R}^5 : \sigma > 0, \lambda \geq 0, \sigma_Y \geq 0\}$ auf

$$\bar{D} = \mathbb{R} \times [\sigma_l, \sigma_u] \times \left(0, \frac{1}{\tau}\right) \times \mathbb{R} \times [\sigma_{Y_l}, \sigma_{Y_u}] \quad (3.15)$$

vor, da dann die Likelihood-Funktion beschränkt ist. Zur numerischen Umsetzung der Maximierung der logarithmierten Likelihood-Funktion (3.14) wählt Honoré eine kompakte Menge $K \subset \mathbb{R}_+$ und nutzt für jeweils fixiertes $k \in K$ die Parameterelimination $\sigma_Y^2 = k\sigma^2$ und löst das Problem für die vier Parameter μ, σ, λ und μ_Y . Anschließend wird in [41] dasjenige $k \in K$ gewählt, für das die logarithmierte Likelihood-Funktion ihr absolutes Maximum annimmt. Man erhält auf diese Weise eine konsistente Schätzung $\underline{p}^\delta \in \bar{D} \subset D$ als Lösung des allgemeinen inversen Problems für empirische Returndaten. Durch die Beschränkung des Definitionsbereichs gelingt es die logarithmierte Maximum-Likelihood-Funktion numerisch zu maximieren. Die Inkorrektheit des Problems wird damit allerdings noch nicht überwunden. In Kapitel 6 werden wir an verschiedenen numerischen Beispielen zeigen, dass die Maximum-Likelihood-Schätzer der einzelnen Parameter recht instabil sind.

3.2.2 Momentenmethode

In Abschnitt 2.2 haben wir das System (2.19) der Semiinvarianten der Returns mit Lag τ eines Jump-Diffusions-Prozesses in Abhängigkeit der Parameter $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)$ hergeleitet und gesehen, wie man diese in Momente umrechnen kann (vgl. Satz 2.2.8). Daher können wir auch versuchen die Parameter mittels der Momentenmethode aus den ersten k empirischen Momenten $m_{\tau,1}^\delta, m_{\tau,2}^\delta, \dots, m_{\tau,k}^\delta$, die wir aus den empirischen Returns $r_1^\delta, r_2^\delta, \dots, r_n^\delta$ berechnen können, zu schätzen. Da in unserem Fall das System der ersten k Semiinvarianten eine übersichtlichere Struktur hat und wegen Satz 2.2.8 die ersten k Semiinvarianten die gleiche Information wie die ersten k Momente enthalten, nutzten wir die empirischen Semiinvarianten und System (2.19), um die Parameter zu schätzen. Wir wählen also den Parametervektor $\underline{p} \in D$, der Lösung des Least-Squares Problems

$$\sum_{i=1}^k (s_{\tau,i}(\underline{p}) - s_{\tau,i}^\delta)^2 \longrightarrow \min \quad \text{für } \underline{p} \in D$$

ist. Dieses Vorgehen entspricht einfach der verallgemeinerten Momentenmethode (siehe z.B. [34, Chapter 14]). Die empirischen Semiinvarianten $s_{\tau,1}^\delta, s_{\tau,2}^\delta, \dots, s_{\tau,k}^\delta$ berechnen wir dabei mit Hilfe von Lemma 2.2.9 aus den empirischen zentrierten Momenten.

Ein wesentlicher Vorteil dieser Methode ist, dass wir die Momente bzw. Semiinvarianten im Gegensatz zur Dichtefunktion mit endlich vielen Rechenoperationen aus

den Parametern berechnen können. Wir kommen also im Vergleich zur Maximum-Likelihood-Methode ohne eine Approximation aus. Trotzdem vererbt sich die Inkorrektheit des Problems der Parameterschätzung auch auf die Momentenmethode, da nach Bemerkung 2.2.11 das k -te Moment eine unendliche gewichtete Überlagerung der k -ten Momente normalverteilter Zufallsgrößen ist. Die Inkorrektheit verursacht wieder eine Vielzahl numerischer Probleme.

Bei der Betrachtung der (verallgemeinerten) Momentenmethode zur Lösung des inversen Problems der Parameterbestimmung wird der Zusammenhang zum allgemeinen Momentenproblem (siehe z.B. [5]) deutlich. Bei diesem versucht man allgemein aus den Momenten m_1, m_2, \dots eine zugehörige nichtnegative Funktion f mit

$$m_k = \int_a^b x^k f(x) dx \quad (k = 0, 1, \dots)$$

zu bestimmen (vgl. [60, Stichwort: Momentenproblem]). Für das Intervall $(a, b) = (0, \infty)$ wird das Problem auch als Stieltjes Momentenproblem, für $(a, b) = (-\infty, \infty)$ als Hamburger Momentenproblem und für $a < b \in \mathbb{R}$ als Hausdorffsches Momentenproblem bezeichnet. Das Momentenproblem zählt zu den inkorrekten inversen Problemen (vgl. [44]). Die Formulierung

$$m_{\tau,k}(\underline{p}) = \int_{-\infty}^{\infty} x^k f_{r_\tau}(x, \underline{p}) dx \quad (k = 0, 1, \dots)$$

unseres inversen Problems zeigt die Verwandtschaft zum Hamburger Momentenproblem. Wir wissen bereits, dass die gesuchte Funktion f die Dichtefunktion f_{r_τ} der Returns r_τ mit Lag τ eines Jump-Diffusions-Prozesses ist und suchen die fünf Parameter, die diese spezifizieren. Die Inkorrektheit des Momentenproblems vererbt sich aber in gewisser Weise auf unser inverses Problem.

3.3 Analytischer Zugang

Unser Ziel ist es eine approximative Lösung \underline{p}^δ von \underline{p}^* zu bestimmen, die stabil von den empirischen Returns $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ der Trajektorie S^δ des Zufallprozesses (2.7) abhängt. Wir können von r^δ die empirische Dichtefunktion berechnen.

Bemerkung 3.3.1 Im Weiteren werden wir bei der Betrachtung der empirischen Dichtefunktion von r^* bzw. r^δ die Zerlegung \mathcal{I} (vgl. Definition 3.1.1) nicht explizit angeben. Wir nehmen an, dass sie geeignet gewählt ist.

Falls wir exakte Daten S^* für alle Zeitpunkte $t \in T = [0, \infty)$ zur Parameteridentifikation zur Verfügung hätten, würde r^* für ein beliebig gewähltes Lag τ unendlich viele Realisierungen der Zufallsgröße r_τ bereitstellen. In diesem Fall würde dann die zum Lag τ berechnete empirische Dichtefunktion (3.1) wegen Satz 3.1.2 der Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ entsprechen. Somit gibt es einen eindeutigen Zusammenhang zwischen

einer von dem Parametervektor $\underline{p} \in D$ abhängigen Trajektorie und der Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$.

Wir nutzen den oben erläuterten Zusammenhang zwischen S^* und der Dichtefunktion zur Parameteridentifikation und führen den folgenden Operator des direkten Problems ein.

Definition 3.3.2 (Vorwärtsoperator) Der nichtlineare Operator $A : \underline{p} \mapsto f_{r_\tau}$ des direkten Problems (Vorwärtsoperator) bildet den Parametervektor $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in D = \{\underline{p} \in \mathbb{R}^5 : \sigma > 0, \lambda \geq 0, \sigma_Y \geq 0\}$ auf die Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ ab, der die Reihendarstellung

$$[A(\underline{p})](x) = f_{r_\tau}(x, \underline{p}) = \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) \quad (x \in \mathbb{R}) \quad (3.16)$$

zugrunde liegt.

Daher lösen wir zur Bestimmung des Parametervektors aus einer Trajektorie S^* das folgende inverse Problem, wobei sich f^* , wie oben beschrieben, aus S^* ergibt.

Definition 3.3.3 (Inverses Problem – IP) Aus der (exakten) Dichtefunktion f^* des Returns r_τ mit Lag τ ist der Parametervektor $\underline{p}^* \in D$ so zu bestimmen, dass die Beziehung

$$A(\underline{p}^*) = f^* \quad (3.17)$$

erfüllt ist.

1923 führte Hadamard [33] für (bei ihm lineare) Operatorgleichungen

$$F(x) = y, \quad x \in D \subset X, \quad y \in Y$$

zwischen den Banachräumen X und Y den Begriff der Korrektheit ein (siehe z.B. [38]).

Definition 3.3.4 Die Operatorgleichung heißt korrekt nach Hadamard, falls die folgenden drei Bedingungen erfüllt sind.

1. F ist surjektiv, d.h. zu jedem $y \in Y$ existiert ein $x \in X$ mit $F(x) = y$ (Existenzbedingung).
2. F ist injektiv, d.h. falls für $y \in F(D) = \{y \in Y : y = F(x), x \in D\}$ die Beziehung $y = F(x_1) = F(x_2)$ mit $x_1, x_2 \in D$ erfüllt ist, folgt $x_1 = x_2$ (Eindeutigkeitsbedingung).
3. F ist stetig invertierbar, d.h. falls der inverse Operator F^{-1} existiert, so ist er stetig. Aus der Konvergenz $y_n \rightarrow y_0$ in Y folgt dann die Konvergenz der Urbilder $F^{-1}(y_n) \rightarrow F^{-1}(y_0)$ in X (Stabilitätsbedingung).

Ist mindestens eine der drei Bedingungen nicht erfüllt, wird die Operatorgleichung als inkorrekt bezeichnet.

Bemerkung 3.3.5 Für nichtlineare Operatorgleichungen, zu denen (3.17) zählt, macht die erste Forderung wenig Sinn, da in der Regel selbst die Abschließung des Bildbereichs eine echte Teilmenge von Y ist, d.h. es gilt $\overline{F(D)} \subsetneq Y$. In der inversen Theorie nichtlinearer Probleme (z.B. [23, p. 524, Formel (1.5)]) wird normalerweise die Bedingung der Erreichbarkeit (attainability) gestellt, d.h. für als exakt betrachtete Elemente $y_0 \in Y$ existiert ein $x_0 \in D(F)$ mit $F(x_0) = y_0$.

Wir betrachten nun wieder das inverse Problem aus Definition 3.3.3. Wir identifizieren den Parametervektor $\underline{p} \in D$ anhand der Dichtefunktion, wobei wir die Dichtefunktion aus den Marktdaten bestimmen. Hierfür nutzen wir solche Daten, deren empirischen Eigenschaften die Verwendung des Modells der Jump-Diffusions-Prozesse rechtfertigen. Der stochastische Prozess, der die Trajektorie S der Marktdaten generiert hat, muss natürlich nicht zwingend ein Jump-Diffusions-Prozess sein, sondern kann auch ein stochastischer Prozess mit ähnlicher Charakteristik sein. Deswegen fordern wir an dieser Stelle, wie in Bemerkung 3.3.5 erwähnt, die Bedingung der Erreichbarkeit. Die aus der Trajektorie S abgeleitete empirische Dichtefunktion f^δ der Returns sei eine Approximation der Dichtefunktion f . Dann sichert uns die Forderung der Erreichbarkeit, dass für die Dichtefunktion f ein Parametervektor $\underline{p} \in D$ mit $A(\underline{p}) = f$ existiert.

Da wir jeder vom Zufall ausgewählten Trajektorie die Dichtefunktion f der Returns r_τ zuordnen, haben wir den stochastischen Hintergrund des Problems in die Funktion f transformiert. Somit ist das von uns betrachtete inverse Problem (IP) ein deterministisches Problem. Daher können wir die üblichen Methoden der deterministischen inversen Theorie zur Lösung von (IP) verwenden.

Für den Fall, dass wir die Trajektorie S^* nicht wie im Idealfall angenommen für alle Zeitpunkte $t \in [0, \infty)$ kennen, sondern nur für Zeitpunkte $t \in T \subsetneq (0, \infty]$, liefert die empirische Dichtefunktion $h(r^*)$ nur eine Approximation der Funktion f^* . Daher interpretieren wir $h(r^*)$ als eine mit dem Datenfehlerniveau $\delta > 0$ gegebene Näherung von f^* und bezeichnen diese deshalb mit f^δ , d.h. es gilt mit einem geeignet gewählten Abstandsbegriff $\rho(\cdot, \cdot)$ zwischen zwei Funktionen

$$\rho(f^*, f^\delta) \leq \delta.$$

Nutzen wir empirische Daten r^δ , so erhalten wir aus der empirischen Dichtefunktion $h(r^\delta)$ ebenfalls eine gestörte Funktion von f^* und bezeichnen diese auch mit f^δ .

Bemerkung 3.3.6 Wir wissen, dass die Datenfunktion f^δ eine Approximation der uns unbekannten exakten Funktion f^* ist. Obwohl f^* unbekannt ist, nehmen wir an, dass wir den Fehler $\rho(f^*, f^\delta)$ mit einem uns bekannten $\delta > 0$ abschätzen können. Da sich für jede Datenfunktion f^δ ein anderes $\delta > 0$ ergeben kann, bezeichnet man δ häufig auch als „noise level“. Wir haben bereits motiviert, dass sowohl exakte, aber nur endlich viele, Returndaten $r^* = (r_1^*, \dots, r_n^*)$ als auch leicht gestörte Returndaten $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ auf eine Datenfunktion f^δ führen. Daher werden wir im Weiteren nicht unterscheiden, ob f^δ auf r^* oder r^δ basiert.

Für die Funktion f^δ muss von Gleichung (3.17) keine Lösung existieren. Daher betrachten wir das folgende Extremalproblem

$$\rho(A(\underline{p}), f^\delta) \longrightarrow \min \quad \text{für } \underline{p} \in D,$$

dessen Lösung wir mit $\underline{p}^\delta \in D$ bezeichnen.

Als Abstand ρ zwischen den Funktionen f_1 und f_2 wählen wir den L^2 -Abstand

$$\rho(f_1, f_2) = \sqrt{\int_{\mathbb{R}} (f_1(x) - f_2(x))^2 dx},$$

so dass wir dann speziell das folgende Least-Squares Problem betrachten

$$\Psi(\underline{p}) := \rho(A(\underline{p}), f^\delta)^2 = \|A(\underline{p}) - f^\delta\|_{L^2(\mathbb{R})}^2 \longrightarrow \min \quad \text{für } \underline{p} \in D. \quad (3.18)$$

In den Kapiteln 4 und 5 werden wir die Eigenschaften der Operatorgleichung (3.17) sowie des Extremalproblems (3.18) näher untersuchen.

3.3.1 Modifikation des Abstandsbegriffs

In dem vorherigen Abschnitt haben wir für ρ den L^2 -Abstand gewählt, so dass für die Abweichungen der Dichtefunktion f^\star und der geringfügig gestörten Dichtefunktion f^δ (siehe Bemerkung 3.3.6)

$$\rho(f^\star, f^\delta) = \|f^\star - f^\delta\|_{L^2(\mathbb{R})} \leq \delta \quad (3.19)$$

gilt.

Wir wollen in diesem Abschnitt kurz weitere mögliche Abstandsbegriffe für ρ diskutieren. Neben der L^2 -Norm können wir beispielsweise auch jede beliebige andere L^q -Norm ($1 \leq q \leq \infty$) oder die C -Norm verwenden⁶. Weil wir hier speziell den Abstand zwischen zwei Dichtefunktionen messen wollen, können wir ρ auch von dem Abstand zwischen den zwei zugehörigen Wahrscheinlichkeitsmaßen ableiten.

Da wir in dieser Arbeit ausschließlich den L^2 -Abstand verwenden und hier nur kurz die anderen möglichen Abstandsbegriffe vorstellen, illustrieren wir deren Unterschiede für den Spezialfall der geometrischen Brownschen Bewegung, d.h. die Sprungparameter der Parametervektoren $\underline{p}_1 = (\mu_1, \sigma_1, \lambda_1, \mu_{Y_1}, \sigma_{Y_1})$ und $\underline{p}_2 = (\mu_2, \sigma_2, \lambda_2, \mu_{Y_2}, \sigma_{Y_2})$ erfüllen die Beziehung $\lambda_1(\mu_{Y_1}^2 + \sigma_{Y_1}^2) = \lambda_2(\mu_{Y_2}^2 + \sigma_{Y_2}^2) = 0$. Die Returns r_τ sind dann normalverteilt. Die Abweichungen der zugehörigen Dichtefunktionen

$$f_1(x) = f_{r_\tau}(x, \underline{p}_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) \quad (3.20a)$$

$$f_2(x) = f_{r_\tau}(x, \underline{p}_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right), \quad (3.20b)$$

⁶In Kapitel 4 zeigen wir, dass $f_{r_\tau}(\cdot, \underline{p}) \in L^q(\mathbb{R})$, $1 \leq q \leq \infty$, sowie $f_{r_\tau}(\cdot, \underline{p}) \in C(\mathbb{R})$ für alle $\underline{p} \in D$ gilt.

die wir mit den verschiedenen Abstandsbegriffen messen, lassen sich für diesen Spezialfall explizit berechnen.

Allgemein können wir für Parametervektoren $\underline{p} \in D$ keine expliziten Formeln für die verschiedenen Abstände herleiten. Wir können jedoch die Abstände numerisch berechnen, indem wir die Summation der Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ (siehe (2.15)) nach genügend vielen Gliedern N abbrechen, d.h. die Approximation $f_{r_\tau}^N(\cdot, \underline{p})$ (siehe (3.12)) anstelle von $f_{r_\tau}(\cdot, \underline{p})$ verwenden.

Wir werden nun zunächst zum Vergleich mit den anderen Abstandsbegriffen den L^2 -Abstand zwischen den Dichtefunktionen zweier normalverteilter Zufallsgrößen berechnen.

Beispiel 3.3.7 *Wir betrachten die Dichtefunktionen (3.20a) und (3.20b) und erhalten dann wegen*

$$\begin{aligned} \rho_{L^2}^2(f_1, f_2) &= \int_{\mathbb{R}} (f_1(x) - f_2(x))^2 dx \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\frac{1}{\sigma_1} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) - \frac{1}{\sigma_2} \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right) \right)^2 dx \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{\sigma_1^2} \exp\left(-\frac{(x - \mu_1)^2}{\sigma_1^2}\right) dx + \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{\sigma_2^2} \exp\left(-\frac{(x - \mu_2)^2}{\sigma_2^2}\right) dx \\ &\quad - \frac{1}{2\pi} \int_{\mathbb{R}} \frac{2}{\sigma_1 \sigma_2} \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2} - \frac{(x - \mu_2)^2}{2\sigma_2^2}\right) dx \\ &= \frac{1}{2\sigma_1 \sqrt{\pi}} + \frac{1}{2\sigma_2 \sqrt{\pi}} - \frac{\sqrt{2}}{\sqrt{\pi}(\sigma_1^2 + \sigma_2^2)} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \end{aligned}$$

für den L^2 -Abstand

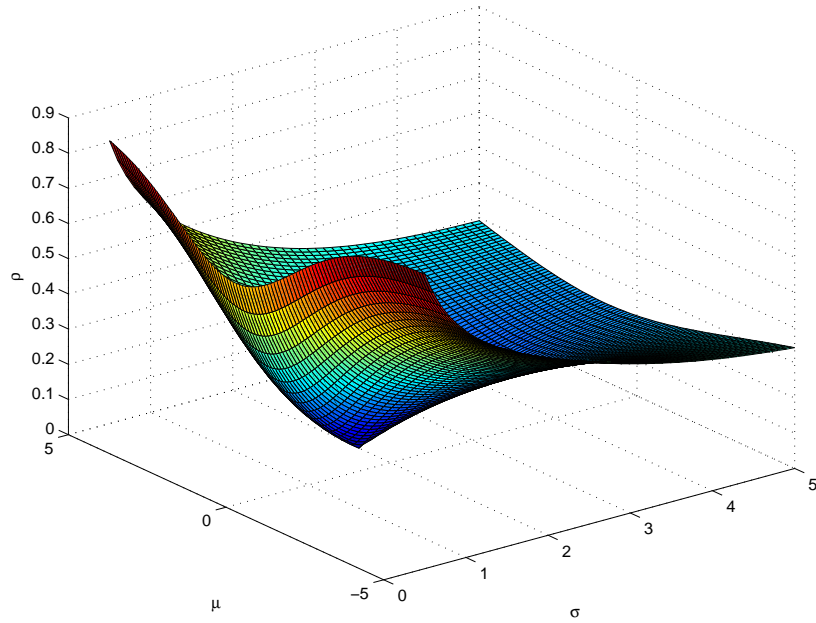
$$\rho_{L^2}(f_1, f_2) = \sqrt{\frac{1}{2\sigma_1 \sqrt{\pi}} + \frac{1}{2\sigma_2 \sqrt{\pi}} - \frac{\sqrt{2}}{\sqrt{\pi}(\sigma_1^2 + \sigma_2^2)} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right)}.$$

In Abbildung 3.6 ist der L^2 -Abstand zwischen den Dichtefunktionen f_1 mit $\mu_1 = 1$ und $\sigma_1 = 2$ sowie der Funktion f_2 in Abhängigkeit der Parameter μ_2 und σ_2 aus Beispiel 3.3.7 zu sehen.

3.3.1.1 Der Hellinger-Abstand

Es seien \mathbb{Q}_1 und \mathbb{Q}_2 zwei Wahrscheinlichkeitsmaße⁷, die bezüglich des Lebesgue Maßes \mathbb{P} absolut stetig sind ($\mathbb{Q} \ll \mathbb{P}$, $\tilde{\mathbb{Q}} \ll \mathbb{P}$). Wegen des Satzes von Radon Nikodym existieren dann die Dichtefunktionen $f_{q_1} = \frac{d\mathbb{Q}_1}{d\mathbb{P}}$ und $f_{q_2} = \frac{d\mathbb{Q}_2}{d\mathbb{P}}$. Somit können wir, basierend auf dem Hellinger-Abstand zwischen den zwei Maßen \mathbb{Q}_1 und \mathbb{Q}_2 (siehe beispielsweise [67,

⁷Die Maße \mathbb{Q}_1 , \mathbb{Q}_2 und \mathbb{P} sind auf dem messbaren Raum (Ω, \mathfrak{F}) definiert.

Abbildung 3.6: L^2 -Abstand

S. 560])

$$\rho_H(Q_1, Q_2) = \sqrt{1 - \int_{\Omega} \sqrt{\frac{dQ}{dP} \frac{d\tilde{Q}}{dP}} dP} = \sqrt{1 - \int_{\mathbb{R}} \sqrt{f_{q_1}(x) f_{q_2}(x)} dx},$$

den Hellinger-Abstand zwischen den Dichtefunktionen f_{q_1} und f_{q_2} einführen.

Definition 3.3.8 Wir definieren mit

$$\rho_H(f_{q_1}, f_{q_2}) = \sqrt{1 - \int_{\mathbb{R}} \sqrt{f_{q_1}(x) f_{q_2}(x)} dx}$$

den Hellinger-Abstand zwischen den Dichtefunktionen f_{q_1} und f_{q_2} .

Aus Definition 3.3.8 folgt unmittelbar $\rho_H(f_{q_1}, f_{q_2}) \leq 1$. Wir können den Hellinger-Abstand mit der Umformung

$$\begin{aligned} 1 - \int_{\mathbb{R}} \sqrt{f_{q_1}(x) f_{q_2}(x)} dx &= \frac{1}{2} \int_{\mathbb{R}} f_{q_1}(x) dx + \frac{1}{2} \int_{\mathbb{R}} f_{q_2}(x) dx - \int_{\mathbb{R}} \sqrt{f_{q_1}(x) f_{q_2}(x)} dx \\ &= \frac{1}{2} \int_{\mathbb{R}} \left(f_{q_1}(x) - 2\sqrt{f_{q_1}(x) f_{q_2}(x)} + f_{q_2}(x) \right) dx \\ &= \frac{1}{2} \int_{\mathbb{R}} \left(\sqrt{f_{q_1}(x)} - \sqrt{f_{q_2}(x)} \right)^2 dx \end{aligned}$$

auch wie folgt berechnen

$$\begin{aligned}\rho_H(f_{q_1}, f_{q_2}) &= \sqrt{1 - \int_{\mathbb{R}} \sqrt{f_{q_1}(x) f_{q_2}(x)} dx} \\ &= \sqrt{\frac{1}{2} \int_{\mathbb{R}} \left(\sqrt{f_{q_1}(x)} - \sqrt{f_{q_2}(x)} \right)^2 dx}. \quad (3.21)\end{aligned}$$

Es lässt sich leicht überprüfen, dass der Hellinger-Abstand für alle Dichtefunktionen f_{q_1} , f_{q_2} und f_{q_3} die Bedingungen

1. $\rho_H(f_{q_1}, f_{q_2}) \geq 0$ sowie $\rho_H(f_{q_1}, f_{q_2}) = 0$, genau dann wenn $f_{q_1} = f_{q_2}$,
2. $\rho_H(f_{q_1}, f_{q_2}) = \rho_H(f_{q_2}, f_{q_1})$ und
3. $\rho_H(f_{q_1}, f_{q_3}) \leq \rho_H(f_{q_1}, f_{q_2}) + \rho_H(f_{q_2}, f_{q_3})$,

und⁸ somit auch die Axiome des metrischen Abstandes erfüllt.

Folgerung 3.3.9 *Mit (3.21) können wir den Hellinger-Abstand in den L^2 -Abstand überführen*

$$\rho_H^2(f_{q_1}, f_{q_2}) = \frac{1}{2} \rho_{L^2}^2(\sqrt{f_{q_1}}, \sqrt{f_{q_2}}).$$

Beispiel 3.3.10 *Wir betrachten wieder die Dichtefunktionen (3.20a) und (3.20b). Dann erhalten wir wegen*

$$\begin{aligned}\rho_H(f_1, f_2)^2 &= 1 - \frac{1}{\sqrt{2\pi\sigma_1\sigma_2}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2} \left(\frac{x^2 - 2x\mu_1 + \mu_1^2}{2\sigma_1^2} + \frac{x^2 - 2x\mu_2 + \mu_2^2}{2\sigma_2^2} \right)\right) dx \\ &= 1 - \frac{1}{\sqrt{\sigma_1\sigma_2} \sqrt{\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{4(\sigma_1^2 + \sigma_2^2)}\right) \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{z^2}{2}\right) dz \\ &= 1 - \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{1}{4} \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}\right)\end{aligned}$$

für den Hellinger-Abstand

$$\rho_H(f_1, f_2) = \sqrt{1 - \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{1}{4} \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}\right)}.$$

Abbildung 3.7 zeigt den Hellinger-Abstand zwischen den Dichtefunktionen f_1 mit $\mu_1 = 1$ und $\sigma_1 = 2$ sowie der Funktion f_2 in Abhängigkeit der Parameter μ_2 und σ_2 aus Beispiel 3.3.10. Zumindest in diesem Beispiel sind keine wesentlichen Unterschiede zum L^2 -Abstand (vgl. Abbildung 3.6) festzustellen.

⁸Mit (3.21) folgt die Dreiecksungleichung für den Hellinger-Abstand unmittelbar aus der Dreiecksungleichung des L^2 -Abstandes.

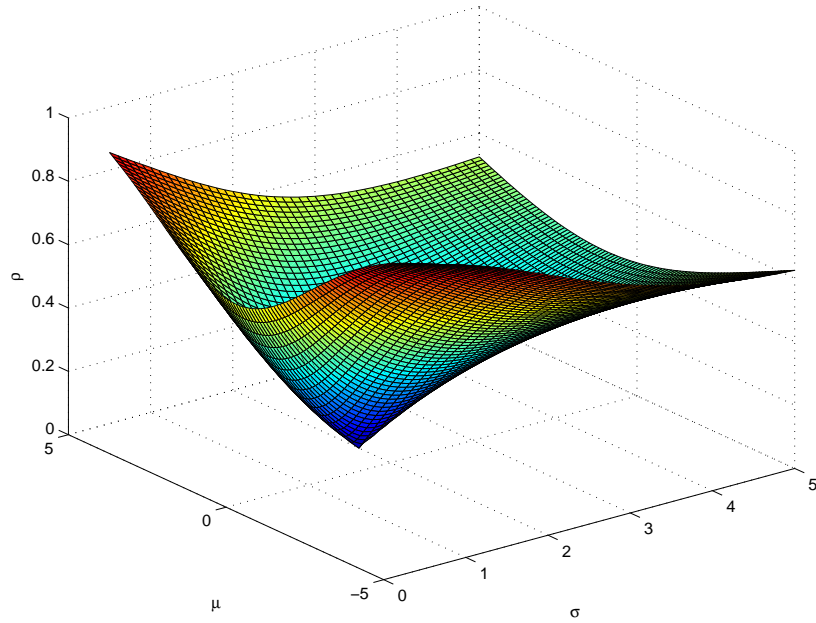


Abbildung 3.7: Hellinger-Abstand

3.3.1.2 Der Variationsabstand

Wir betrachten wieder die zwei Wahrscheinlichkeitsmaße \mathbb{Q}_1 und \mathbb{Q}_2 . Für die zwei Wahrscheinlichkeitsmaße \mathbb{Q}_1 und \mathbb{Q}_2 ist mit

$$\rho_V(\mathbb{Q}_1, \mathbb{Q}_2) = 2 \sup_{A \in \mathfrak{F}} |\mathbb{Q}_1(A) - \mathbb{Q}_2(A)|$$

der Variationsabstand (vgl. [66, S. 360]) gegeben. Falls \mathbb{Q}_1 und \mathbb{Q}_2 bezüglich dem Lebesgue Maß \mathbb{P} absolut stetig sind ($\mathbb{Q} \ll \mathbb{P}$, $\tilde{\mathbb{Q}} \ll \mathbb{P}$), kann man für die Dichtefunktionen $f_{q_1} = \frac{d\mathbb{Q}_1}{d\mathbb{P}}$ und $f_{q_2} = \frac{d\mathbb{Q}_2}{d\mathbb{P}}$ zeigen (siehe [66, III - Corollary 1]), dass die Beziehung

$$\rho_V(\mathbb{Q}_1, \mathbb{Q}_2) = \int_{\mathbb{R}} |f_{q_1}(x) - f_{q_2}(x)| dx =: \rho_V(f_{q_1}, f_{q_2})$$

erfüllt ist. Somit entspricht dann der Variationsabstand der Dichtefunktionen einfach dem L^1 -Abstand.

Bemerkung 3.3.11 Zwischen dem Variationsabstand und dem Hellinger-Abstand besteht folgender Zusammenhang (siehe [66, III §9 - Theorem 1])

$$2\rho_H^2(f_{q_1}, f_{q_2}) \leq \rho_V(f_{q_1}, f_{q_2}) \leq \sqrt{8}\rho_H(f_{q_1}, f_{q_2}). \quad (3.22)$$

Beispiel 3.3.12 Wir wollen nun für die Dichtefunktionen (3.20a) und (3.20b) den Variationsabstand berechnen. Falls $\sigma_1 \neq \sigma_2$ besitzen die Dichtefunktionen f_1 und f_2 die

zwei Schnittpunkte

$$x_1 = \frac{\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2} - \sqrt{\left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}\right)^2 - \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) \left(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2} + \ln \frac{\sigma_1^2}{\sigma_2^2}\right)}}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}}$$

$$x_2 = \frac{\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2} + \sqrt{\left(\frac{\mu_1}{\sigma_1^2} - \frac{\mu_2}{\sigma_2^2}\right)^2 - \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}\right) \left(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2} + \ln \frac{\sigma_1^2}{\sigma_2^2}\right)}}{\frac{1}{\sigma_1^2} - \frac{1}{\sigma_2^2}},$$

andernfalls haben die zwei Funktionen für $\sigma_1 = \sigma_2$ einen Schnittpunkt in

$$x_1 = \frac{\mu_1 + \mu_2}{2}.$$

In diesem Fall setzen wir dann $x_2 = x_1$ und erhalten

$$\begin{aligned} \rho_V(f_1, f_2) &= \int_{-\infty}^{\infty} |f_1(x) - f_2(x)| dx \\ &= \left| \int_{-\infty}^{x_1} (f_1(x) - f_2(x)) dx \right| + \left| \int_{x_1}^{x_2} (f_1(x) - f_2(x)) dx \right| + \left| \int_{x_2}^{\infty} (f_1(x) - f_2(x)) dx \right| \\ &= \begin{cases} 2 \left| 2\Phi\left(\frac{\mu_2 - \mu_1}{2\sigma_1}\right) - 1 \right| & \text{für } \sigma_1 = \sigma_2 \\ \left| \Phi\left(\frac{x_1 - \mu_1}{\sigma_1}\right) - \Phi\left(\frac{x_1 - \mu_2}{\sigma_2}\right) \right| \\ + \left| \Phi\left(\frac{x_2 - \mu_2}{\sigma_2}\right) - \Phi\left(\frac{x_2 - \mu_1}{\sigma_1}\right) \right| \\ + \left| \Phi\left(\frac{x_2 - \mu_1}{\sigma_1}\right) - \Phi\left(\frac{x_1 - \mu_1}{\sigma_1}\right) \right| \\ - \left| \Phi\left(\frac{x_2 - \mu_2}{\sigma_2}\right) + \Phi\left(\frac{x_1 - \mu_2}{\sigma_2}\right) \right| & \text{für } \sigma_1 \neq \sigma_2. \end{cases} \end{aligned}$$

In Abbildung 3.8 ist der Variationsabstand zwischen den Dichtefunktionen f_1 mit $\mu_1 = 1$ und $\sigma_1 = 2$ sowie der Funktion f_2 in Abhängigkeit der Parameter μ_2 und σ_2 aus Beispiel 3.3.12 abgebildet. Auch hier können wir kein grundsätzlich anderes Verhalten im Vergleich zum L^2 -Abstand (vgl. Abbildung 3.6) beobachten.

3.3.1.3 Der Kullback-Leibler-Abstand

Als letztes Beispiel für eine mögliche Abstandswahl wollen wir noch den Kullback-Leibler-Abstand betrachten, der eng mit dem Informationsabstand und der Maximum-Entropie (siehe z.B. [21], [22], [47]) verbunden ist.

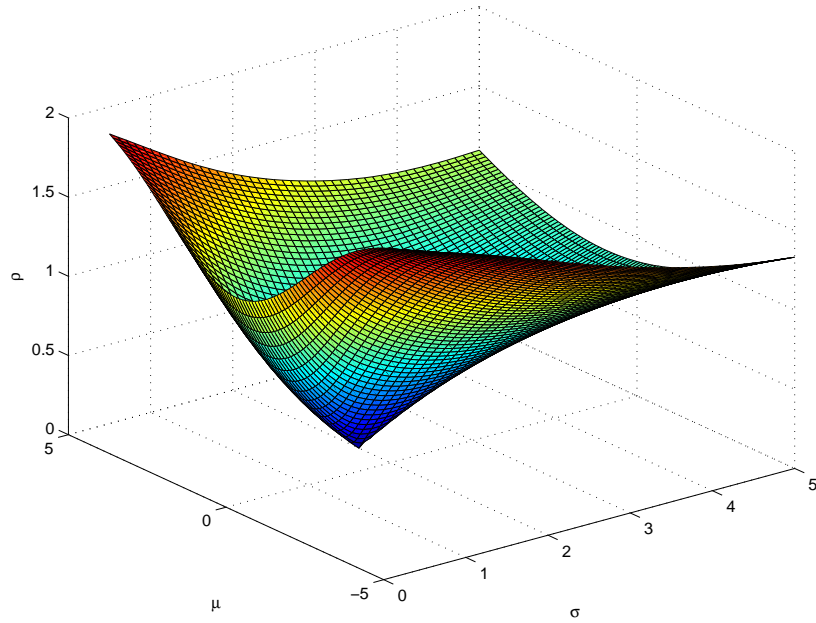


Abbildung 3.8: Variationsabstand

Definition 3.3.13 Der Kullback-Leibler-Abstand zwischen den Dichtefunktionen f_{q_1} und f_{q_2} ($f_{q_2}(x) > 0, \forall x \in \mathbb{R}$) ist durch⁹

$$\rho_{KL}(f_{q_1}, f_{q_2}) = \int_{\mathbb{R}} f_{q_1}(x) \ln \frac{f_{q_1}(x)}{f_{q_2}(x)} dx \quad (3.23)$$

gegeben.

Obwohl der Kullback-Leibler-Abstand weder die Symmetrieforderung noch die Dreiecksungleichung erfüllt, wird er häufig verwendet, um den Informationsabstand zweier Dichtefunktionen zu messen. Man kann durch die Modifikation $\tilde{\rho}_{KL}(f_{q_1}, f_{q_2}) = \frac{1}{2}(\rho_{KL}(f_{q_1}, f_{q_2}) + \rho_{KL}(f_{q_2}, f_{q_1}))$ erreichen (siehe [73]), dass die Symmetrieforderung eines metrischen Abstandbegriffes erfüllt ist. Wir werden hier zur Illustration den Kullback-Leibler-Abstand aus Definition 3.3.13 verwenden.

Beispiel 3.3.14 Wir erhalten mit den Dichtefunktionen (3.20a) und (3.20b) für den Kullback-Leibler-Abstand

$$\begin{aligned} \rho_{KL}(f_1, f_2) &= \int_{-\infty}^{\infty} f_1(x) \ln \frac{f_1(x)}{f_2(x)} dx \\ &= \int_{-\infty}^{\infty} f_1(x) \ln \left(\frac{\frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right)}{\frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right)} \right) dx \end{aligned}$$

⁹Falls die zu den Dichtefunktionen f_{q_1} bzw. f_{q_2} gehörenden Wahrscheinlichkeitsmaße äquivalent sind, deren Nullmengen also übereinstimmen, beeinflussen diese Nullmengen nicht das Integral (3.23), so dass wir dann nicht zwingend $f_{q_2}(x) > 0$ fordern müssen.

$$\begin{aligned}
&= \left(\ln \frac{\sigma_2}{\sigma_1} - \frac{\sigma_2^2 \mu_1^2 - \sigma_1^2 \mu_2^2}{2\sigma_1^2 \sigma_2^2} \right) \int_{-\infty}^{\infty} f_1(x) dx + \frac{\mu_1 \sigma_2^2 - \mu_2 \sigma_1^2}{\sigma_1^2 \sigma_2^2} \int_{-\infty}^{\infty} x f_1(x) dx \\
&\quad - \frac{\sigma_2^2 - \sigma_1^2}{2\sigma_1^2 \sigma_2^2} \int_{-\infty}^{\infty} x^2 f_1(x) dx \\
&= \ln \frac{\sigma_2}{\sigma_1} - \frac{\sigma_2^2 \mu_1^2 - \sigma_1^2 \mu_2^2}{2\sigma_1^2 \sigma_2^2} + \frac{\mu_1 \sigma_2^2 - \mu_2 \sigma_1^2}{\sigma_1^2 \sigma_2^2} \mu_1 - \frac{\sigma_2^2 - \sigma_1^2}{2\sigma_1^2 \sigma_2^2} (\mu_1^2 + \sigma_1^2) \\
&= \ln \frac{\sigma_2}{\sigma_1} - \frac{1}{2} + \frac{\sigma_1^2}{2\sigma_2^2} + \frac{(\mu_1 - \mu_2)^2}{2\sigma_2^2}.
\end{aligned}$$

In Abbildung 3.9, die den Kullback-Leibler-Abstand zwischen den Dichtefunktionen f_1 mit $\mu_1 = 1$ und $\sigma_1 = 2$ sowie der Funktion f_2 in Abhängigkeit der Parameter μ_2 und σ_2 zeigt, ist ein deutlicher Unterschied zu den anderen betrachteten Abständen zu sehen. Die zwei Querschnitte in Abbildung 3.10 verdeutlichen diesen Unterschied nochmals. In der linken Graphik haben wir die verschiedenen Abstände der Dichtefunktionen f_1 und f_2 für festes $\sigma_2 = 1$ in Abhängigkeit von μ_2 und in der rechten Graphik für fixiertes $\mu_2 = 0$ in Abhängigkeit von σ_2 dargestellt.

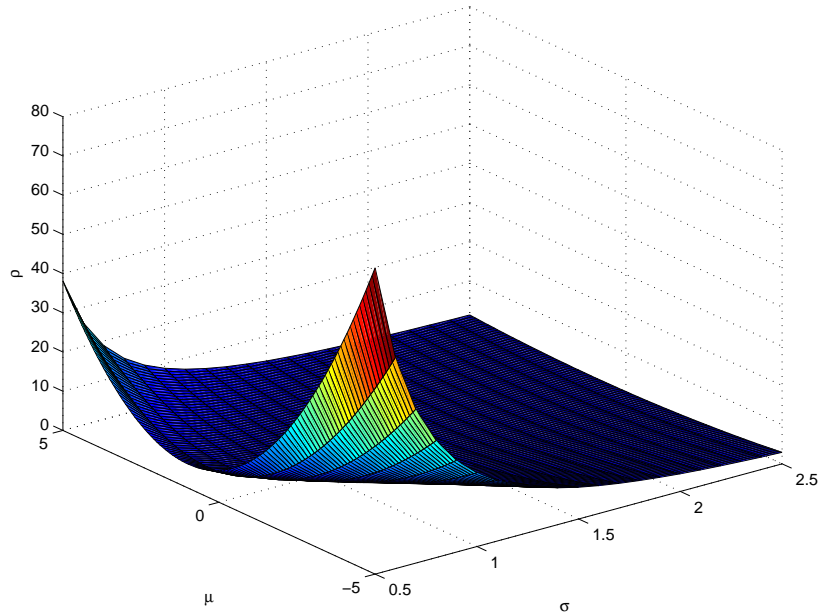


Abbildung 3.9: Kullback-Leibler-Abstand

Anhand der kleinen Fallstudie, die wir zu den verschiedenen Abständen mittels der Dichtefunktionen (3.20a) und (3.20b) durchgeführt haben, konnten wir keine wesentlichen Unterschiede zwischen dem L^2 -Abstand, dem Hellinger-Abstand und dem Variationsabstand (L^1 -Abstand) feststellen. Wegen Bemerkung 3.3.11 und Folgerung 3.3.9 gibt es auch im Allgemeinen einen Zusammenhang zwischen diesen Abstandsbegriffen. Daher sind bei der Verwendung des Hellinger-Abstandes bzw. Variationsabstandes anstelle des L^2 -Abstandes in (3.18) bzw. (3.19) keine signifikanten Veränderungen zu

erwarten. Deutliche Abweichungen traten dagegen bei der Verwendung des Kullback-Leibler-Abstandes auf, so dass bei dessen Verwendung in (3.18) bzw. (3.19) anstelle des L^2 -Abstandes einige Auswirkungen zu erwarten sind. Allerdings erfüllt dieser Abstands begriff nicht die Axiome einer Metrik.

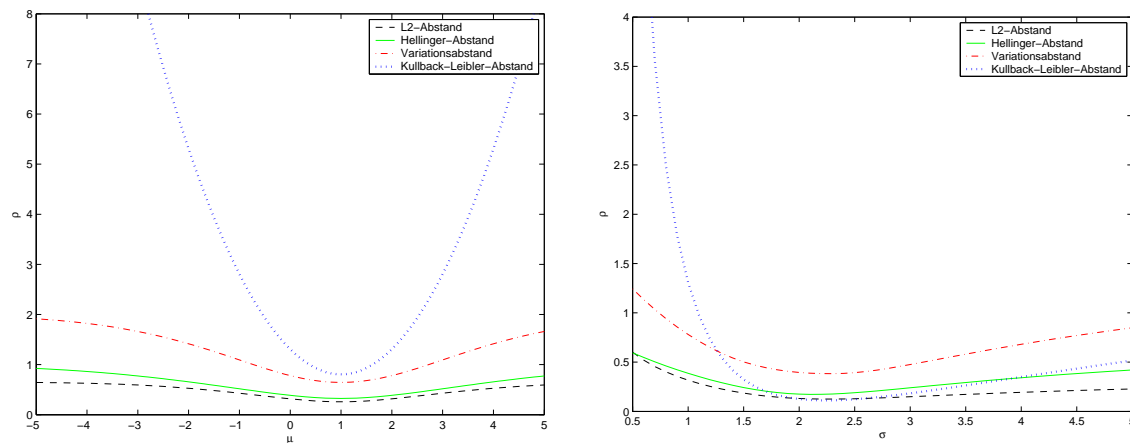


Abbildung 3.10: Querschnitte der verschiedenen Abstände zwischen f_1 mit $\mu_1 = 1$ sowie $\sigma_1 = 2$ und f_2 in Abhängigkeit von μ_2 bzw. σ_2

Kapitel 4

Eigenschaften des Vorwärtsoperators

In Kapitel 3 haben wir die Operatorgleichung (3.17) des inversen Problems der Parameteridentifizierung und in Definition 3.3.4 sowie Bemerkung 3.3.5 die Begriffe Korrektheit bzw. Inkorrekttheit eingeführt. Wir werden nun mit Hinblick auf die Korrektheit bzw. Inkorrekttheit der Operatorgleichung (3.17) die Eigenschaften des Vorwärtsoperators untersuchen. Bisher haben wir immer zu einem festen Parameter $\underline{p} \in D$ die zugehörige Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ betrachtet. Daher müssen wir also zunächst überprüfen, welchen Einfluss die Variation der Parameter aus dem Definitionsbereich (2.4) auf die Funktion $f_{r_\tau} : (x, \underline{p}) \rightarrow \mathbb{R}_+$ hat.

Zunächst ist wegen der Nichtnegativität aller Reihenglieder und

$$\begin{aligned} [A(\underline{p})](x) &= \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) \\ &\leq \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \frac{1}{\sqrt{2\pi}} \\ &\leq \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau}} \frac{1}{\sqrt{2\pi}} = \frac{1}{\sigma\sqrt{2\pi\tau}} < \infty \quad (\forall \sigma > 0) \end{aligned} \tag{4.1}$$

der Vorwärtsoperator A für alle $\underline{p} \in D$ wohldefiniert. Die Funktion $[A(\underline{p})](x)$ ist für feste \underline{p} eine unendliche Überlagerung gewichteter Dichtefunktionen $g_j(x, \underline{p})$ normalverteilter Zufallsgrößen mit Erwartungswert $\tilde{\mu}\tau + j\mu_Y$ und Varianz $\sigma^2\tau + j\sigma_Y^2$, d.h.

$$g_j(x, \underline{p}) = \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right), \tag{4.2}$$

wobei die Gewichte $w_j = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!}$ der Wahrscheinlichkeit, dass genau j Sprünge auftreten, entsprechen. Solch eine Überlagerung so genannter Glockenkurven tritt auch bei

der Betrachtung von linearen Fredholmschen Integraloperatoren

$$[F(x)](s) = \int_a^b k(s, t)x(t) dt = y(s) \quad (4.3)$$

mit speziellen Kernfunktionen $k \in C([c, d] \times [a, b])$ auf. Im Beispiel 2.3 in [38] stellt die Kernfunktion $k(s, \cdot)$ beispielsweise für jedes $s \in [c, d]$ eine Glockenkurve dar. In den Banachräumen $X = C([a, b])$ und $Y = C([c, d])$ ist die Operatorgleichung (4.3) kompakt und somit inkorrekt. Unsere Operatorgleichung (3.17) bildet zwar von dem endlichdimensionalen Raum $D \subset \mathbb{R}^5$ ab und wird anstelle der Integralgleichung durch eine unendliche Summe beschrieben, die Überlagerung der Glockenkurven hat aber einen ähnlichen glättenden Charakter, der die Inkorrektheitsphänomene verursacht.

Der Beweis von Satz 4.1.1 wird zeigen, dass die Funktion $[A(\underline{p})](x)$ für jeden festen Parametervektor $\underline{p} \in D$ auf \mathbb{R} gleichmäßig stetig ist. Somit können wir im Folgenden die Integrations- und Summationsreihenfolge vertauschen und erhalten

$$\begin{aligned} \int_{\mathbb{R}} [A(\underline{p})](x) dx &= \int_{\mathbb{R}} \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) dx \\ &= \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} \int_{\mathbb{R}} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) dx \\ &= \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} = e^{-\lambda\tau} e^{\lambda\tau} = 1. \end{aligned}$$

Für festes $\underline{p} \in D$ ist die Funktion $[A(\underline{p})](x)$ also wirklich eine Dichtefunktion. Wegen $f_{r_\tau}(x, \underline{p}) \leq \frac{1}{\sqrt{\sigma^2\tau}} < \infty$ ($\forall \sigma > 0$), $f_{r_\tau}(x, \underline{p}) \geq 0$ und $\int_{\mathbb{R}} f_{r_\tau}(x, \underline{p}) dx = 1$ folgt unmittelbar $f_{r_\tau}(\cdot, \underline{p}) \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ und somit $f_{r_\tau}(\cdot, \underline{p}) \in L^q(\mathbb{R})$ $1 \leq q \leq \infty$.

Zur stabilen Lösung des inversen Problems benötigen wir die Stetigkeit des Vorwärtsoperators (siehe auch [17]).

4.1 Stetigkeit des Operators

Für einen festen Parametervektor $\underline{p} \in D$ ist die Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ über \mathbb{R} stetig. Bisher haben wir allerdings noch nicht die Auswirkung variabler Parameter auf die Funktion $f_{r_\tau} : (x, \underline{p}) \mapsto y \in \mathbb{R}_+$ analysiert. Daher untersuchen wir zunächst die Stetigkeit dieser Funktion.

Die Reihenglieder

$$f_j(x, \underline{p}) = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) \quad (4.4)$$

der Funktion (3.16) sind alle stetig bezüglich $(x, \underline{p}) \in \mathbb{R} \times D$. Außerdem konvergiert die Reihe (3.16) für alle $x \in \mathbb{R}$ und $\underline{p} \in D$. Trotzdem genügt die Stetigkeit aller Reihenglieder (4.4) und die Konvergenz der Reihe (3.16) nicht, um sofort auf die Stetigkeit der Grenzfunktion $f_{r_\tau}(x, \underline{p})$ zu schlussfolgern. Die gleichmäßige Konvergenz der Reihe gegen die Grenzfunktion ist aber für die Stetigkeit der Grenzfunktion hinreichend.

Theorem 4.1.1 *Die Funktion $f_{r_\tau} : (x, \underline{p}) \mapsto \mathbb{R}$ ist für alle $(x, \underline{p}) \in \mathbb{R} \times D$ stetig.*

Beweis. Um die Stetigkeit der Grenzfunktion $f_{r_\tau}(x, \underline{p})$ der Reihe (3.16) zu untersuchen, betrachten wir zunächst die absoluten Reihenglieder (4.4). Diese können wir wegen $e^{-\lambda\tau} \leq 1$ und $\frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \leq \frac{1}{\sqrt{\sigma^2\tau}}$ mit

$$|f_j(x, \underline{p})| = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} \frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right) \leq \frac{(\lambda\tau)^j}{j! \sqrt{2\pi\tau}\sigma} \quad (4.5)$$

für alle $(x, \underline{p}) \in \mathbb{R} \times D$ abschätzen.

Für jeden Parametervektor $\underline{p}_0 \in D$ existiert ein $\delta_0 > 0$, so dass für die δ_0 -Umgebung $B_{\delta_0}(\underline{p}_0) = \{\underline{p} \in D : \|\underline{p} - \underline{p}_0\| \leq \delta_0\} \subset D$ die Beziehung

$$\sigma_{\min} := \sigma_{\min}(B_{\delta_0}(\underline{p}_0)) = \min_{\sigma: \underline{p}=(\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in B_{\delta_0}(\underline{p}_0)} \sigma > 0 \quad (4.6)$$

erfüllt ist. In dieser Umgebung existieren für jeden einzelnen Parameter des Parametervektors $\underline{p} \in B_{\delta_0} \subset D$ Obergrenzen sowie Untergrenzen. Insbesondere gibt es also ein

$$\lambda_{\max} := \max_{\lambda: \underline{p}=(\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in B_{\delta_0}(\underline{p}_0)} \lambda < \infty$$

und nach (4.6) ein σ_{\min} , so dass $\lambda \leq \lambda_{\max} < \infty$ und $0 < \sigma_{\min} \leq \sigma$ für alle Parametervektoren $\underline{p} \in B_{\delta_0}(\underline{p}_0) \subset D$ gilt. Deswegen können wir für alle $\underline{p} \in B_{\delta_0}(\underline{p}_0) \subset D$ wegen (4.5) die Reihenglieder durch

$$|f_j(x, \underline{p})| \leq \frac{(\lambda_{\max}\tau)^j}{j! \sqrt{2\pi\tau}\sigma_{\min}} =: c_j$$

abschätzen. Die Majorantenreihe $\{c_j\}_{j=1}^\infty$ der Reihe (3.16) ist wegen

$$\sum_{j=0}^\infty c_j = \sum_{j=0}^\infty \frac{(\lambda_{\max}\tau)^j}{j! \sqrt{2\pi\tau}\sigma_{\min}} = \frac{e^{\lambda_{\max}\tau}}{\sqrt{2\pi\tau}\sigma_{\min}} < \infty$$

konvergent. Daher folgt aus dem Weierstraßschen Majorantenkriterium (siehe z.B. [28, S. 443]) die gleichmäßige Konvergenz der Reihe (3.16) gegen die Grenzfunktion $f_{r_\tau}(x, \underline{p})$ für alle $(x, \underline{p}) \in \mathbb{R} \times B_{\delta_0}(\underline{p}_0)$.

Die Stetigkeit der einzelnen Reihenglieder und die gleichmäßige Konvergenz sichern uns nun die Stetigkeit der Grenzfunktion, was wir nun als Nächstes zeigen. Diese Aussage findet man zum Beispiel in Satz 28.7 in [40, S. 162].

Wir bezeichnen mit $p_n(x, \underline{p})$ die n -te Partialsumme

$$p_n(x, \underline{p}) = \sum_{j=0}^n f_j(x, \underline{p}) \quad n = 0, 1, \dots$$

und $u_n(x, \underline{p})$ den Rest

$$u_n(x, \underline{p}) = \sum_{j=n+1}^{\infty} f_j(x, \underline{p}) .$$

Aus der gleichmäßigen Konvergenz der Reihe (3.16) für alle $(x, \underline{p}) \in \mathbb{R} \times B_{\delta_0}(\underline{p}_0)$ folgt unmittelbar, dass für ein beliebiges $\varepsilon > 0$ ein $\hat{n} = \hat{n}(\varepsilon, \delta_0)$ mit $u_n(x, \underline{p}) < \frac{\varepsilon}{3}$ für alle $(x, \underline{p}) \in \mathbb{R} \times B_{\delta_0}(\underline{p}_0)$ und $n \geq \hat{n}$ existiert.

Wir betrachten nun diejenigen $(x, \underline{p}) \in B_{\delta_1}(x_0) \times B_{\delta_2}(\underline{p}_0) = \{(x, \underline{p}) \in \mathbb{R} \times D : |x - x_0| \leq \delta_1, \|\underline{p} - \underline{p}_0\| \leq \delta_2\}$, die sich in einer Umgebung um (x_0, \underline{p}_0) befinden, wobei $\delta_2 \leq \delta_0$ gilt. Da jede Partialsumme $p_n(x, \underline{p})$ als endliche Summe stetiger Funktionen $f_j(x, \underline{p})$ bezüglich x und \underline{p} stetig ist, können wir für \hat{n} und hinreichend kleine $\delta_1 = \delta_1(\varepsilon)$ und $\delta_2 = \delta_2(\varepsilon)$ für alle $(x, \underline{p}) \in B_{\delta_1}(x_0) \times B_{\delta_2}(\underline{p}_0)$ den Abstand der Partialsummen $|p_{\hat{n}}(x, \underline{p}) - p_{\hat{n}}(x_0, \underline{p}_0)|$ ebenfalls mit $\frac{\varepsilon}{3}$ abschätzen und erhalten somit die Abschätzung

$$\begin{aligned} |f(x, \underline{p}) - f(x_0, \underline{p}_0)| &= |p_{\hat{n}}(x, \underline{p}) + u_{\hat{n}}(x, \underline{p}) - p_{\hat{n}}(x_0, \underline{p}_0) - u_{\hat{n}}(x_0, \underline{p}_0)| \\ &\leq |p_{\hat{n}}(x, \underline{p}) - p_{\hat{n}}(x_0, \underline{p}_0)| + |u_{\hat{n}}(x, \underline{p})| + |u_{\hat{n}}(x_0, \underline{p}_0)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon . \end{aligned}$$

Aus der gleichmäßigen Konvergenz der Reihe und der Stetigkeit der Summanden folgt also die Stetigkeit bezüglich x und \underline{p} der Grenzfunktion $f_{r_\tau}(x, \underline{p})$ für alle $(x, \underline{p}) \in \mathbb{R} \times B_{\delta_0}(\underline{p}_0)$. Insbesondere ist $f_{r_\tau}(x, \underline{p})$ stetig in (x, \underline{p}_0) für alle $x \in \mathbb{R}$. Da $\underline{p}_0 \in D$ beliebig gewählt war, ist $f_{r_\tau}(x, \underline{p})$ in $\mathbb{R} \times D$ stetig \square

Bemerkung 4.1.2 Mit dem Beweis von Theorem 4.1.1 erhalten wir sogar für jede beliebige Konstante $C \in \mathbb{R}_+$ und für alle (x, \underline{p}) aus der kompakten Menge $[-C, C] \times B_{\delta_0}(\underline{p}_0)$ die gleichmäßige Stetigkeit der Funktion $f_{r_\tau}(x, \underline{p})$.

Aus der Stetigkeit der Funktion $f_{r_\tau} : (x, \underline{p}) \mapsto y \in \mathbb{R}_+$ folgt nun die Stetigkeit des Operators bezüglich der L^2 -Norm. Wir zeigen zunächst die Stetigkeit in der C-Norm.

Theorem 4.1.3 *Der Operator A aus (3.16) ist für alle $\underline{p} \in D$ bezüglich der C-Norm stetig.*

Beweis. Der Operator A bildet den Parametervektor $\underline{p} \in D$ auf die Dichtefunktion $x \mapsto f_{r_\tau}(x, \underline{p})$ ab, d.h.

$$A(\underline{p}) = f_{r_\tau}(\cdot, \underline{p}) .$$

Wir betrachten die Parametervektoren \underline{p} , die sich in einer δ -Umgebung des Parametervektors \underline{p}_0 befinden, d.h. die Parametervektoren $\underline{p} \in B_\delta(\underline{p}_0) \subset D$. Um die Stetigkeit bezüglich der C-Norm nachzuweisen, müssen wir

$$\left\| A(\underline{p}) - A(\underline{p}_0) \right\|_{C(\mathbb{R})} = \left\| f_{r_\tau}(\cdot, \underline{p}) - f_{r_\tau}(\cdot, \underline{p}_0) \right\|_{C(\mathbb{R})} = \sup_{x \in \mathbb{R}} \left| f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0) \right| < \varepsilon \quad (4.7)$$

für hinreichend kleines $\delta = \delta(\varepsilon)$ zeigen.

Nach Bemerkung 4.1.2 folgt, dass für eine beliebige Konstante C die Grenzfunktion $f_{r_\tau}(x, \underline{p})$ für alle $(x, \underline{p}) \in [-C, C] \times B_\delta(\underline{p}_0)$ gleichmäßig stetig ist. Daraus folgt, dass für hinreichend kleines $\delta = \delta(\varepsilon)$

$$\left| f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0) \right| < \varepsilon \quad \forall \underline{p} \in B_\delta(\underline{p}_0), \quad \forall x \in [-C, C], \quad (4.8)$$

gilt. Um die C-Norm in (4.7) abzuschätzen, benötigen wir noch das Verhalten der Dichtefunktion außerhalb des Intervalls $[-C, C]$.

Die Dichtefunktion (3.16) ist eine unendliche Überlagerung gewichteter Glockenkurven. Da für jeden einzelnen Summanden der Reihe die jeweiligen Funktionswerte für betragsmäßig wachsende x exponentiell abklingen, gehen die Funktionswerte der Dichtefunktion für große x sehr schnell gegen null. Wir zeigen nun, dass für jedes $\varepsilon > 0$ ein C mit

$$f_{r_\tau}(x, \underline{p}) < \frac{\varepsilon}{2} \quad \forall (x, \underline{p}) \in (\mathbb{R} \setminus [-C, C]) \times B_\delta(\underline{p}_0) \quad (4.9)$$

existiert.

Um (4.9) nachzuweisen, unterscheiden wir die folgenden zwei Fälle. Zunächst betrachten wir den einfachen Fall $\sigma_{Y_{\min}} = \min_{\underline{p} \in B_\delta(\underline{p}_0)} \sigma_Y > 0$, d.h. in der Umgebung von \underline{p}_0 befinden sich nur Parameter, die rein zufällige Sprunghöhen implizieren. Die ungewichteten Glockenkurven (4.2) können wir in diesem Fall für alle $(x, \underline{p}) \in \mathbb{R} \times B_\delta(\underline{p}_0)$ mit

$$g_j(x, \underline{p}) = \frac{1}{\sqrt{\sigma^2 \tau + j \sigma_Y^2}} \phi \left(\frac{x - (\tilde{\mu} \tau + j \mu_Y)}{\sqrt{\sigma^2 \tau + j \sigma_Y^2}} \right) \leq \frac{1}{\sqrt{2\pi} \sqrt{\sigma_{\min}^2 \tau + j \sigma_{Y_{\min}}^2}} \quad (4.10)$$

abschätzen. Die obere Grenze

$$\tilde{g}_j := \frac{1}{\sqrt{2\pi} \sqrt{\sigma_{\min}^2 \tau + j \sigma_{Y_{\min}}^2}}$$

ist bezüglich j (streng) monoton fallend und strebt für j gegen unendlich gegen null. Deswegen existiert ein $\hat{j} = \hat{j}(\varepsilon)$, so dass $g_j(x, \underline{p}) < \frac{\varepsilon}{2}$ für alle $j \geq \hat{j}$ und $(x, \underline{p}) \in \mathbb{R} \times B_\delta(\underline{p}_0)$ gilt. In diesem Fall können wir sogar explizit die Form von \hat{j}

$$\hat{j}(\varepsilon) = \left\lceil \frac{1}{\sigma_{Y_{\min}}^2} \left(\frac{2}{\pi \varepsilon^2} - \sigma_{\min}^2 \tau \right) \right\rceil + 1$$

angeben. Nun bestimmen wir für alle $j < \hat{j}$ jeweils eine Grenze $C_j = C_j(\varepsilon, \underline{p})$, so dass $g_j(x, \underline{p}) < \frac{\varepsilon}{2}$ für alle $(x, \underline{p}) \in (\mathbb{R} \setminus [-C_j, C_j]) \times B_\delta(\underline{p}_0)$ gilt. Für jedes $j < \hat{j}$ existiert ein

solches endliches C_j , da die Glockenkurven $g_j(x, \underline{p})$ exponentiell abklingen. Wir können in diesem Fall C_j sogar explizit angeben. Da für einen festen Parametervektor \underline{p}

$$g_j(x, \underline{p}) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2\tau + j\sigma_Y^2}} \exp\left(-\frac{(x - \tilde{\mu}\tau - j\mu_Y)^2}{2(\sigma^2\tau + j\sigma_Y^2)}\right) < \frac{\varepsilon}{2}$$

genau dann erfüllt ist, wenn

$$x > \sqrt{(\sigma^2\tau + j\sigma_Y^2) \ln \frac{2}{\varepsilon^2\pi(\sigma^2\tau + j\sigma_Y^2)}} + \tilde{\mu}\tau + j\mu_Y$$

oder

$$x < -\sqrt{(\sigma^2\tau + j\sigma_Y^2) \ln \frac{2}{\varepsilon^2\pi(\sigma^2\tau + j\sigma_Y^2)}} + \tilde{\mu}\tau + j\mu_Y$$

gilt und für jeden einzelnen Parameter μ , σ , λ , μ_Y und σ_Y Ober- und Untergrenzen $\mu_{\min} < \mu < \mu_{\max}$, usw. für alle $\underline{p} \in B_\delta(\underline{p}_0)$ existieren, können wir C_j wie folgt berechnen

$$C_j(\varepsilon, \underline{p}) = \max \left\{ \left| \sqrt{(\sigma_{\max}^2\tau + j\sigma_{Y_{\max}}^2) \ln \frac{2}{\varepsilon^2\pi(\sigma_{\min}^2\tau + j\sigma_{Y_{\min}}^2)}} + \tilde{\mu}_{\max} + j\mu_{Y_{\max}} \right|, \right. \\ \left. \left| -\sqrt{(\sigma_{\max}^2\tau + j\sigma_{Y_{\max}}^2) \ln \frac{2}{\varepsilon^2\pi(\sigma_{\min}^2\tau + j\sigma_{Y_{\min}}^2)}} + \tilde{\mu}_{\min} + j\mu_{Y_{\min}} \right| \right\}.$$

Damit können wir für den ersten Fall $\sigma_{Y_{\min}} > 0$ nun C als die größte Grenze, d.h. $C = \max\{C_1, \dots, C_{j-1}\}$, wählen und erhalten für alle $(x, \underline{p}) \in (\mathbb{R} \setminus [-C, C]) \times B_\delta(\underline{p}_0)$ wie in (4.9) gefordert

$$f_{r_\tau}(x, \underline{p}) = \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} g_j(x, \underline{p}) < \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} \frac{\varepsilon}{2} = \frac{\varepsilon}{2} e^{-\lambda\tau} e^{\lambda\tau} = \frac{\varepsilon}{2}.$$

Nun betrachten wir den zweiten Fall $\sigma_{Y_{\min}} = \min_{\underline{p} \in B_\delta(\underline{p}_0)} \sigma_Y = 0$. Da die Abschätzung \tilde{g}_j der Glockenkurven in (4.10) nun wegen $\sigma_{Y_{\min}} = 0$ unabhängig von j ist, benötigen wir in diesem Fall eine andere Argumentation. Wir können die einzelnen Reihenglieder (4.4) der Dichtefunktion für $(x, \underline{p}) \in \mathbb{R} \times B_\delta(\underline{p}_0)$ mit

$$f_j(x, \underline{p}) = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!\sqrt{\sigma^2\tau}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau}}\right) \leq e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!\sqrt{2\pi}\sqrt{\sigma_{\min}^2\tau}} =: \tilde{f}_j(\lambda)$$

abschätzen. Die obere Grenze $\tilde{f}_j(\lambda)$ konvergiert für j gegen unendlich für alle $\underline{p} \in B_\delta(\underline{p}_0)$ sehr schnell gegen null, so dass für alle $0 < \tilde{\varepsilon} < 1$ wegen der Äquivalenz der Beziehungen

$$\tilde{f}_j(\lambda) = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!\sqrt{2\pi}\sqrt{\sigma_{\min}^2\tau}} < \tilde{\varepsilon}^{j+1} \quad (4.11)$$

und

$$\frac{1}{\tilde{\varepsilon}} \frac{1}{j!} \left(\frac{\lambda\tau}{\tilde{\varepsilon}} \right)^j = \frac{(\lambda\tau)^j}{j! \tilde{\varepsilon}^{j+1}} < e^{\lambda\tau} \sqrt{2\pi\sigma_{\min}^2\tau} \quad (4.12)$$

ein endliches $\hat{j} = \hat{j}(\tilde{\varepsilon})$ existiert mit $f_j(x, \underline{p}) < \tilde{\varepsilon}^{j+1}$ für alle $(x, \underline{p}) \in \mathbb{R} \times B_\delta(\underline{p}_0)$ und $j \geq \hat{j}(\tilde{\varepsilon})$. Da für alle festen $\tilde{\varepsilon} > 0$ und λ mit $0 \leq \lambda_{\min} \leq \lambda \leq \lambda_{\max}$ die linke Seite von (4.12) $\frac{(\lambda\tau)^j}{j! \tilde{\varepsilon}^{j+1}} = \frac{1}{\tilde{\varepsilon}} \frac{1}{j!} \left(\frac{\lambda\tau}{\tilde{\varepsilon}} \right)^j$ für $j \rightarrow \infty$ gegen null strebt, existiert $\hat{j}(\tilde{\varepsilon})$ wirklich, so dass (4.11) für alle $j \geq \hat{j}(\tilde{\varepsilon})$ und $(x, \underline{p}) \in \mathbb{R} \times B_\delta(\underline{p}_0)$ erfüllt ist.

Damit existieren wieder für alle $j < \hat{j}$ Grenzen C_j , so dass außerhalb des Intervalls $[-C_j, C_j]$ alle Funktionswerte der Funktion $f_j(x, \underline{p})$ kleiner als $\tilde{\varepsilon}^{j+1}$ für alle Parametervektoren aus der δ -Umgebung $B_\delta(\underline{p}_0)$ von \underline{p}_0 sind. Da die Funktionen $f_j(x, \underline{p})$ mit $w_j = e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!}$ gewichtete Glockenkurven sind, klingen diese exponentiell schnell ab. Deswegen existiert für jedes $j < \hat{j}$ ein solches C_j und wir wählen auch in diesem Fall $C = \max \{C_1, \dots, C_{\hat{j}-1}\}$.

Somit können wir nun mit $\tilde{\varepsilon} = \frac{\varepsilon}{\varepsilon+2}$ wegen $f_j(x, \underline{p}) < \tilde{\varepsilon}^{j+1}$ die Dichtefunktion auch im zweiten Fall $\sigma_{Y_{\min}} = 0$ für alle $(x, \underline{p}) \in (\mathbb{R} \setminus [-C, C]) \times B_\delta(\underline{p}_0)$, wie in (4.9) beschrieben, mit

$$f_{r_\tau}(x, \underline{p}) = \sum_{j=0}^{\infty} f_j(x, \underline{p}) < \sum_{j=0}^{\infty} \tilde{\varepsilon}^{j+1} = \tilde{\varepsilon} \frac{1}{1 - \tilde{\varepsilon}} = \frac{\varepsilon}{2}$$

abschätzen. Diese Argumentation gilt auch für den ersten Fall $\sigma_{\min} > 0$, das dortige Vorgehen liefert aber explizite Abschätzungen für \hat{j} und C .

Wir können nun also wegen (4.8) und (4.9) den Abstand der Dichte $f_{r_\tau}(\cdot, \underline{p})$ zur Dichte $f_{r_\tau}(\cdot, \underline{p}_0)$ für alle Parametervektoren $\underline{p} \in B_\delta(\underline{p}_0)$ bezüglich der C -Norm wie in (4.7) gefordert abschätzen. Hierfür betrachten wir den Ausdruck

$$\begin{aligned} \|A(\underline{p}) - A(\underline{p}_0)\|_{C(\mathbb{R})} &= \|f_{r_\tau}(\cdot, \underline{p}) - f_{r_\tau}(\cdot, \underline{p}_0)\|_{C(\mathbb{R})} = \sup_{x \in \mathbb{R}} |f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)| \\ &= \max \left\{ \sup_{x \in \mathbb{R} \setminus [-C, C]} |f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)|, \max_{x \in [-C, C]} |f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)| \right\}. \end{aligned} \quad (4.13)$$

Den ersten Teil aus (4.13) können wir wegen $f_{r_\tau}(x, \underline{p}) < \frac{\varepsilon}{2}$ (vgl. (4.9)) und

$$|f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)| \leq |f_{r_\tau}(x, \underline{p})| + |f_{r_\tau}(x, \underline{p}_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

mit ε abschätzen. Für den zweiten Teil aus (4.13) erhalten wir wegen

$$|f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)| < \varepsilon$$

(vgl. (4.8)) die gleiche Abschätzung, so dass

$$\|A(\underline{p}) - A(\underline{p}_0)\|_{C(\mathbb{R})} = \|f_{r_\tau}(\cdot, \underline{p}) - f_{r_\tau}(\cdot, \underline{p}_0)\|_{C(\mathbb{R})} = \sup_{x \in \mathbb{R}} |f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)| < \varepsilon$$

gilt und damit die Stetigkeit des Operators in der C -Norm nachgewiesen ist \square

Theorem 4.1.4 *Der Operator A aus (3.16) ist für alle $\underline{p} \in D$ bezüglich der L^q -Norm für alle $q > 1$ stetig.*

Beweis. Aus der Stetigkeit von A bezüglich der C -Norm folgt unmittelbar, dass für alle $\underline{p} \in B_\delta(\underline{p}_0)$

$$\sup \left| f(x, \underline{p}) - f(x, \underline{p}_0) \right| < \varepsilon$$

gilt. Außerdem haben wir

$$\int_{\mathbb{R}} \left| f(x, \underline{p}) - f(x, \underline{p}_0) \right| dx \leq \int_{\mathbb{R}} \left| f(x, \underline{p}) \right| + \left| f(x, \underline{p}_0) \right| dx \leq 2,$$

so dass

$$\begin{aligned} \left\| A(\underline{p}) - A(\underline{p}_0) \right\|_{L^q}^q &= \int_{\mathbb{R}} \left| f(x, \underline{p}) - f(x, \underline{p}_0) \right|^q dx \\ &\leq \left\{ \sup \left| f(x, \underline{p}) - f(x, \underline{p}_0) \right| \right\}^{q-1} \int_{\mathbb{R}} \left| f(x, \underline{p}) - f(x, \underline{p}_0) \right| dx \\ &< 2\varepsilon^{q-1} =: \tilde{\varepsilon}^q \end{aligned}$$

und somit für alle $1 < q < \infty$ die Stetigkeit von A bezüglich der L^q -Norm folgt \square

Wir haben nun die Stetigkeit des Operators bezüglich der C -Norm und L^q -Norm ($1 < q < \infty$) gezeigt. In Abschnitt 3.3.1 haben wir noch weitere Normen bzw. Abstandsbe-
griffe diskutiert. Wir zeigen nun zunächst die Stetigkeit bezüglich der Hellinger-Norm

$$\left\| f_{r_\tau}(\cdot, \underline{p}) - f_{r_\tau}(\cdot, \underline{p}_0) \right\|_H = \rho_H \left(f_{r_\tau}(\cdot, \underline{p}), f_{r_\tau}(\cdot, \underline{p}_0) \right).$$

Theorem 4.1.5 *Der Operator A aus (3.16) ist für alle $\underline{p} \in D$ bezüglich der Hellinger-Norm stetig.*

Beweis. Wegen (3.21) gilt für die Hellinger-Norm

$$\begin{aligned} \left\| A(\underline{p}) - A(\underline{p}_0) \right\|_H &= \rho_H \left(f_{r_\tau}(\cdot, \underline{p}), f_{r_\tau}(\cdot, \underline{p}_0) \right) = \sqrt{1 - \int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p}) f_{r_\tau}(x, \underline{p}_0)} dx} \\ &= \sqrt{\frac{1}{2} \int_{\mathbb{R}} \left(\sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right)^2 dx}. \end{aligned}$$

Um die Stetigkeit bezüglich der Hellinger-Norm zu zeigen, müssen wir für Parametervektoren \underline{p} , die sich in einer δ -Umgebung des Parametervektors \underline{p}_0 befinden, d.h. $\underline{p} \in B_\delta(\underline{p}_0) \subset D$, für hinreichend kleines $\delta = \delta(\varepsilon)$ die Beziehung

$$\left\| A(\underline{p}) - A(\underline{p}_0) \right\|_H = \sqrt{\frac{1}{2} \int_{\mathbb{R}} \left(\sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right)^2 dx} < \varepsilon$$

nachweisen.

Wegen der Stetigkeit des Operators A bzgl. der C -Norm haben wir für $\underline{p} \in B_\delta(\underline{p}_0) \subset D$ (vgl. (4.7))

$$\|A(\underline{p}) - A(\underline{p}_0)\|_{C(\mathbb{R})} = \sup_{x \in \mathbb{R}} |f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)| < \hat{\varepsilon}. \quad (4.14)$$

Wir zeigen nun, dass für alle $x \in \mathbb{R}$ und alle $\underline{p} \in B_\delta(\underline{p}_0)$ sowie $\underline{p}_0 \in D$ die Abschätzung

$$\left| \sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right| < \tilde{\varepsilon} \quad (4.15)$$

gilt. Dazu betrachten wir $z_1 > 0$ und $z_2 > 0$, wobei ohne Beschränkung der Allgemeinheit $z_1 > z_2$ sei¹. Für $z_1 \geq \varepsilon_1$ erhalten wir mit $L = \frac{1}{\sqrt{\varepsilon_1}}$ wegen

$$\frac{|\sqrt{z_1} - \sqrt{z_2}|}{|z_1 - z_2|} = \left| \frac{\sqrt{z_1} - \sqrt{z_2}}{(\sqrt{z_1} - \sqrt{z_2})(\sqrt{z_1} + \sqrt{z_2})} \right| = \frac{1}{\sqrt{z_1} + \sqrt{z_2}} < \frac{1}{\sqrt{z_1}} \leq \frac{1}{\sqrt{\varepsilon_1}} = L$$

die Beziehung

$$|\sqrt{z_1} - \sqrt{z_2}| < L |z_1 - z_2|. \quad (4.16)$$

Für $z_1 < \varepsilon_1$ und $z_2 < \varepsilon_1$ erhalten wir

$$|\sqrt{z_1} - \sqrt{z_2}| \leq \sqrt{z_1} + \sqrt{z_2} < 2\sqrt{\varepsilon_1}. \quad (4.17)$$

Damit bekommen wir nun falls $\max \left\{ \sqrt{f_{r_\tau}(x, \underline{p})}, \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right\} \geq \varepsilon_1$ wegen (4.14) und (4.16) die folgende Abschätzung

$$\left| \sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right| < L |f_{r_\tau}(x, \underline{p}) - f_{r_\tau}(x, \underline{p}_0)| < \frac{1}{\sqrt{\varepsilon_1}} \hat{\varepsilon}.$$

Andernfalls erhalten wir für $\sqrt{f_{r_\tau}(x, \underline{p})} < \varepsilon_1$ und $\sqrt{f_{r_\tau}(x, \underline{p}_0)} < \varepsilon_1$ mit (4.17) unmittelbar

$$\left| \sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right| < 2\sqrt{\varepsilon_1}.$$

Für die Wahl $\varepsilon_1 = \hat{\varepsilon}$ ergibt sich somit für alle $x \in \mathbb{R}$ und alle $\underline{p} \in B_\delta(\underline{p}_0)$ sowie $\underline{p}_0 \in D$

$$\left| \sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right| < 2\sqrt{\hat{\varepsilon}} =: \tilde{\varepsilon}.$$

Damit können wir nun für alle $\underline{p} \in B_\delta(\underline{p}_0)$ die Hellinger-Norm mit der Konstante $C < \infty$ wie folgt abschätzen

$$\begin{aligned} \|A(\underline{p}) - A(\underline{p}_0)\|_H^2 &= \int_{\mathbb{R}} \frac{1}{2} \left(\sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right)^2 dx \\ &\leq \frac{1}{2} \sup_{x \in \mathbb{R}} \left| \sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right| \int_{\mathbb{R}} \left| \sqrt{f_{r_\tau}(x, \underline{p})} - \sqrt{f_{r_\tau}(x, \underline{p}_0)} \right| dx \\ &\leq \frac{\tilde{\varepsilon}}{2} \left(\int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p})} dx + \int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p}_0)} dx \right) \leq \frac{\tilde{\varepsilon}}{2} C =: \varepsilon^2. \end{aligned}$$

¹Falls $z_2 = 0$ erhalten wir für $|z_1| = |z_1 - z_2| < \varepsilon_1$ unmittelbar $\sqrt{z_1} = |\sqrt{z_1} - \sqrt{z_2}| < \sqrt{\varepsilon_1}$.

Wir müssen also nur noch zeigen, dass eine endliche Konstante C mit

$$\int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p})} dx + \int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p}_0)} dx \leq C \quad (4.18)$$

existiert. Hierfür betrachten wir für $\underline{p} \in B_\delta(\underline{p}_0)$ mit

$$\begin{aligned} \lambda_{\max} &= \max_{\lambda: \underline{p}=(\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in B_\delta(\underline{p}_0)} \lambda < \infty, \\ \sigma_{\max} &= \max_{\sigma: \underline{p}=(\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in B_\delta(\underline{p}_0)} \sigma < \infty \quad \text{und} \\ \sigma_{Y \max} &= \max_{\sigma_Y: \underline{p}=(\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in B_\delta(\underline{p}_0)} \sigma_Y < \infty \end{aligned}$$

die Abschätzung

$$\begin{aligned} \int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p})} dx &= \int_{\mathbb{R}} \sqrt{\sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right)} dx \\ &\leq \int_{\mathbb{R}} \sum_{j=0}^{\infty} \sqrt{e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right)} dx \quad (4.19) \\ &\leq \sum_{j=0}^{\infty} \left(\sqrt{\frac{(\lambda\tau)^j}{j!}} \int_{\mathbb{R}} \sqrt{\frac{1}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right)} dx \right) \\ &= \sqrt{2} (2\pi)^{\frac{1}{4}} \sum_{j=0}^{\infty} \left(\sqrt{\frac{(\lambda\tau)^j}{j!}} (\sigma^2\tau + j\sigma_Y^2)^{\frac{1}{4}} \right) \\ &\leq \sqrt{2} (2\pi)^{\frac{1}{4}} \sum_{j=0}^{\infty} \left(\sqrt{\frac{(\lambda_{\max}\tau)^j}{j!}} (\sigma_{\max}^2\tau + j\sigma_{Y \max}^2)^{\frac{1}{4}} \right). \end{aligned}$$

Mit dem Quotientenkriterium können wir zeigen, dass die Reihe $\sum_{j=0}^{\infty} q_j$ mit

$$q_j := \sqrt{\frac{(\lambda_{\max}\tau)^j}{j!}} (\sigma_{\max}^2\tau + j\sigma_{Y \max}^2)^{\frac{1}{4}}$$

in der letzten Zeile von (4.19) wegen

$$\lim_{j \rightarrow \infty} \frac{q_{j+1}}{q_j} = \lim_{j \rightarrow \infty} \sqrt{\frac{\lambda_{\max}\tau}{j+1}} \left(\frac{\sigma_{\max}^2\tau + (j+1)\sigma_{Y \max}^2}{\sigma_{\max}^2\tau + j\sigma_{Y \max}^2} \right)^{\frac{1}{4}} = 0 < 1$$

für alle $\underline{p} \in B_\delta(\underline{p}_0)$ konvergiert. Somit existiert für diese Parametervektoren eine Konstante $\tilde{C} < \infty$ mit

$$\sum_{j=0}^{\infty} \left(\sqrt{\frac{(\lambda_{\max}\tau)^j}{j!}} (\sigma_{\max}^2\tau + j\sigma_{Y \max}^2)^{\frac{1}{4}} \right) \leq \tilde{C},$$

so dass wir schließlich für alle $\underline{p} \in B_\delta(\underline{p}_0)$

$$\int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p})} dx + \int_{\mathbb{R}} \sqrt{f_{r_\tau}(x, \underline{p}_0)} dx \leq 2\sqrt{2} (2\pi)^{\frac{1}{4}} \tilde{C} =: C < \infty$$

erhalten, womit obige Aussage bewiesen ist \square

Mit der Stetigkeit bezüglich der Hellinger-Norm können wir nun die Stetigkeit bezüglich der L^1 -Norm zeigen.

Theorem 4.1.6 *Der Operator A aus (3.16) ist für alle $\underline{p} \in D$ bezüglich der L^1 -Norm stetig.*

Beweis. In Abschnitt 3.3.1.2 haben wir gezeigt, dass der Variationsabstand zwischen den Dichtefunktionen zweier absolut stetiger Zufallsgrößen dem L^1 -Abstand entspricht. Mit (3.22) erhalten wir dann

$$\begin{aligned} \|A(\underline{p}) - A(\underline{p}_0)\|_{L^1(\mathbb{R})} &= \|f_{r_\tau}(\cdot, \underline{p}) - f_{r_\tau}(\cdot, \underline{p}_0)\|_{L^1(\mathbb{R})} \\ &= \rho_V(f_{r_\tau}(\cdot, \underline{p}), f_{r_\tau}(\cdot, \underline{p}_0)) \\ &\leq \sqrt{8} \rho_H(f_{r_\tau}(\cdot, \underline{p}), f_{r_\tau}(\cdot, \underline{p}_0)) \\ &= \sqrt{8} \|A(\underline{p}) - A(\underline{p}_0)\|_H, \end{aligned}$$

so dass aus der Stetigkeit des Operators bezüglich der Hellinger-Norm (Theorem 4.1.5) unmittelbar die Stetigkeit bezüglich der L^1 -Norm folgt \square

Bemerkung 4.1.7 Wir können Theorem 4.1.6 auch ohne Verwendung von Beziehung (3.22) beweisen, da für Parametervektoren $\underline{p} \in B_\delta(\underline{p}_0)$, wegen der Stetigkeit des Operators bezüglich der Hellinger-Norm, aus (4.15) und (4.18) (vgl. Beweis von Theorem 4.1.5) die Abschätzung

$$\begin{aligned} \|A(\underline{p}) - A(\underline{p}_0)\|_{L^1(\mathbb{R})} &= \int_{\mathbb{R}} |f_{r_\tau}(\cdot, \underline{p}) - f_{r_\tau}(\cdot, \underline{p}_0)| dx \\ &= \int_{\mathbb{R}} \left| \left(\sqrt{f_{r_\tau}(\cdot, \underline{p})} - \sqrt{f_{r_\tau}(\cdot, \underline{p}_0)} \right) \left(\sqrt{f_{r_\tau}(\cdot, \underline{p})} + \sqrt{f_{r_\tau}(\cdot, \underline{p}_0)} \right) \right| dx \\ &\leq \tilde{\varepsilon} \int_{\mathbb{R}} \left(\sqrt{f_{r_\tau}(\cdot, \underline{p})} + \sqrt{f_{r_\tau}(\cdot, \underline{p}_0)} \right) dx \\ &\leq \tilde{\varepsilon} C =: \varepsilon \end{aligned}$$

und somit die Stetigkeit bezüglich der L^1 -Norm folgt.

Insgesamt haben wir nun die Stetigkeit des Operators bezüglich der C -Norm, der L^q -Norm ($1 \leq q < \infty$) sowie der Hellinger-Norm gezeigt. Der in Abschnitt 3.3.1 eingeführte Kullback-Leibler-Abstand erfüllt nicht die metrischen Axiome. Daher untersuchen wir auch nicht die Stetigkeit des Operators A bezüglich dieses Abstandes.

Zur Bestimmung einer Lösung des inversen Problems (IP) spielen nach Definition 3.3.4 und Bemerkung 3.3.5 deren Eindeutigkeit und Stabilität eine zentrale Rolle. Die Frage der Eindeutigkeit einer Lösung $\underline{p}^* \in D$ zu gegebener Dichtefunktion f^* von Gleichung (3.17) hängt unmittelbar mit der Injektivität der Operatorgleichung auf D zusammen.

4.2 Injektivität des Operators und Stabilitätsbetrachtungen

Wir wollen in diesem Abschnitt die Korrektheit der Operatorgleichung

$$A(\underline{p}) = f \quad \underline{p} \in D \subset \mathbb{R}^5, \quad f \in Y \quad (4.20)$$

mit dem nichtlinearen Operator $A : X = \mathbb{R}^5 \rightarrow Y$ im Raumpaars (X, Y) näher untersuchen. In Abschnitt 4.1 haben wir die Stetigkeit des Vorwärtsoperators für die Bildräume $Y = L^q(\mathbb{R})$ ($1 < q < \infty$), $Y = C(\mathbb{R})$ und $Y = H$ mit der entsprechenden Norm $\|\cdot\|_Y$ nachgewiesen.

Wir untersuchen zunächst die Injektivität des Operators A mit

$$[A(\underline{p})](x) = f_{\tau\tau}(x, \underline{p}) = \sum_{j=0}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j! \sqrt{\sigma^2\tau + j\sigma_Y^2}} \phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right)$$

für Parametervektoren aus dem Definitionsbereich (2.4), wobei wegen (2.3) und (2.6)

$$\tilde{\mu} = \mu - \frac{\sigma^2}{2} - \lambda \left(e^{\mu_Y + \frac{1}{2}\sigma_Y^2} - 1 \right)$$

gilt.

Lemma 4.2.1 *Der Operator A ist auf D nicht injektiv.*

Beweis. Da die Parametervektoren $\underline{p}_1 = (\mu, \sigma, \lambda, 0, 0)^T$ und $\underline{p}_2 = (\mu, \sigma, 0, \mu_Y, \sigma_Y)^T$ beide auf die gleiche Dichtefunktion

$$[A(\underline{p}_1)](x) = [A(\underline{p}_2)](x) = \frac{1}{\sqrt{\sigma^2\tau}} \phi\left(\frac{x - \tilde{\mu}\tau}{\sqrt{\sigma^2\tau}}\right) \quad (x \in \mathbb{R}), \quad (4.21)$$

abgebildet werden, ist der Operator A offensichtlich nicht injektiv auf D □

Bemerkung 4.2.2 Die in (4.21) auftretende Funktion ist eine Normalverteilungsdichte, da im Fall des ersten Parametervektors \underline{p}_1 die Sprunghöhe fast sicher null beträgt und für \underline{p}_2 fast sicher keine Sprünge mehr auftreten.

Allerdings sind die trivialen Fälle $\underline{p}_1 = (\mu, \sigma, \lambda, 0, 0)$ und $\underline{p}_2 = (\mu, \sigma, 0, \mu_Y, \sigma_Y)$ mit $A(\underline{p}_1) = A(\underline{p}_2)$ die einzigen Beispiele für die Nichtinjektivität. Das folgende Theorem liefert die Injektivität des Operators bezüglich des eingeschränkten Definitionsbereichs und wurde von Starkloff et al. in [69] bewiesen. Wir wiederholen den Beweis im Folgenden zum Verständnis der Zusammenhänge.

Theorem 4.2.3 *Der Operator A ist auf dem eingeschränkten Definitionsbereich $\hat{D} = \{\underline{p} \in D : \lambda(\sigma_Y^2 + \mu_Y^2) \neq 0\}$ injektiv.*

Beweis. Wir betrachten zwei Parametervektoren

$$\underline{p}_1 = (\mu_1, \sigma_1, \lambda_1, \mu_{Y1}, \sigma_{Y1}) \in \hat{D} \quad \text{und} \quad \underline{p}_2 = (\mu_2, \sigma_2, \lambda_2, \mu_{Y2}, \sigma_{Y2}) \in \hat{D},$$

für die die Beziehung

$$A(\underline{p}_1) = A(\underline{p}_2) \tag{4.22}$$

erfüllt ist. Aus (4.22) folgt unmittelbar, dass die charakteristischen Funktionen übereinstimmen müssen, da diese eindeutig von der Wahrscheinlichkeitsverteilung bestimmt werden. Somit stimmen auch die logarithmierten charakteristischen Funktionen überein, d.h. für alle $\theta \in \mathbb{R}$ gilt $\ln \varphi_{r_\tau}(\theta, \underline{p}_1) = \ln \varphi_{r_\tau}(\theta, \underline{p}_2)$. Im Weiteren sei ohne Beschränkung der Allgemeinheit $\tau = 1$, so dass für alle $\theta \in \mathbb{R}$ die Beziehung

$$\begin{aligned} i\tilde{\mu}_1\theta - \frac{\sigma_1^2}{2}\theta^2 + \lambda_1 \left(\exp \left(-\frac{\sigma_{Y1}^2}{2}\theta^2 + i\mu_{Y1}\theta \right) - 1 \right) \\ = i\tilde{\mu}_2\theta - \frac{\sigma_2^2}{2}\theta^2 + \lambda_2 \left(\exp \left(-\frac{\sigma_{Y2}^2}{2}\theta^2 + i\mu_{Y2}\theta \right) - 1 \right) \end{aligned} \tag{4.23}$$

erfüllt sein muss. Dabei haben wir wegen (2.3) und (2.6)

$$\tilde{\mu}_1 = \mu_1 - \frac{\sigma_1^2}{2} - \lambda_1 \left(e^{\mu_{Y1} + \frac{1}{2}\sigma_{Y1}^2} - 1 \right) \quad \text{und} \quad \tilde{\mu}_2 = \mu_2 - \frac{\sigma_2^2}{2} - \lambda_2 \left(e^{\mu_{Y2} + \frac{1}{2}\sigma_{Y2}^2} - 1 \right).$$

Sowohl auf der linken, als auch auf der rechten Seite der Gleichung (4.23) stehen bezüglich θ unendlich oft differenzierbare komplexwertige Funktionen. Nach dreimaliger Ableitung beider Seiten in (4.23) erhalten wir mit den Polynomen

$$\begin{aligned} P_1(\theta) &:= \lambda_1 (-\sigma_{Y1}^2\theta + i\mu_{Y1}) \left((-\sigma_{Y1}^2\theta + i\mu_{Y1})^2 - 3\sigma_{Y1}^2 \right) \\ &= \lambda_1 (-\sigma_{Y1}^6\theta^3 + 3i\mu_{Y1}\sigma_{Y1}^4\theta^2 + (3\sigma_{Y1}^4 + 3\mu_{Y1}^2\sigma_{Y1}^2)\theta - 3i\sigma_{Y1}^2\mu_{Y1} - i\mu_{Y1}^3) \\ P_2(\theta) &:= \lambda_2 (-\sigma_{Y2}^2\theta + i\mu_{Y2}) \left((-\sigma_{Y2}^2\theta + i\mu_{Y2})^2 - 3\sigma_{Y2}^2 \right) \\ &= \lambda_2 (-\sigma_{Y2}^6\theta^3 + 3i\mu_{Y2}\sigma_{Y2}^4\theta^2 + (3\sigma_{Y2}^4 + 3\mu_{Y2}^2\sigma_{Y2}^2)\theta - 3i\sigma_{Y2}^2\mu_{Y2} - i\mu_{Y2}^3) \\ Q_1(\theta) &:= -\frac{\sigma_{Y1}^2}{2}\theta^2 + i\mu_{Y1}\theta \\ Q_2(\theta) &:= -\frac{\sigma_{Y2}^2}{2}\theta^2 + i\mu_{Y2}\theta \end{aligned}$$

die Beziehung

$$P_1(\theta) \exp(Q_1(\theta)) = P_2(\theta) \exp(Q_2(\theta)) \quad \forall \theta \in \mathbb{R}. \tag{4.24}$$

Für den Fall $\sigma_{Y_1} = 0$ gilt dann² $P_1(\theta) \equiv -i\lambda_1\mu_{Y_1}^3 \neq 0$. Wegen $|\exp(Q_1(\theta))| > 0$ und $|\exp(Q_2(\theta))| > 0$ kann das Polynom $P_2(\theta)$ ebenfalls keine Nullstellen besitzen, so dass $P_2(\theta)$ folglich eine Konstante sein muss. Dies impliziert $\sigma_{Y_2} = 0$ und $P_2(\theta) \equiv -i\lambda_1\mu_{Y_1}^3$. Somit muss in diesem Fall dann auch $Q_1(\theta) = Q_2(\theta)$ gelten und folglich $\mu_{Y_1} = \mu_{Y_2}$ sowie $\lambda_1 = \lambda_2$.

Seien nun im Weiteren σ_{Y_1} und σ_{Y_2} positiv, so dass P_1 und P_2 für alle $\underline{p}_1 \in \hat{D}$ und $\underline{p}_2 \in \hat{D}$ Polynome dritten Grades³ sind. Wegen $|\exp(Q_1(\theta))| > 0$ und $|\exp(Q_2(\theta))| > 0$ stimmen die drei einfachen Nullstellen von P_1

$$\begin{aligned}\theta_1 &= \frac{i\mu_{Y_1}}{\sigma_{Y_1}^2} \\ \theta_2 &= \frac{i\mu_{Y_1} + \sqrt{3}\sigma_{Y_1}}{\sigma_{Y_1}^2} \\ \theta_3 &= \frac{i\mu_{Y_1} - \sqrt{3}\sigma_{Y_1}}{\sigma_{Y_1}^2}\end{aligned}\tag{4.25}$$

mit denen von P_2 überein. Daher ist für alle $\theta \in \mathbb{R}$ und eine Konstante c die Beziehung $P_1(\theta) = cP_2(\theta)$ erfüllt. Für $\theta = 0$ gilt $\exp(Q_1(0)) = \exp(Q_2(0)) = 1$, woraus sofort $P_1(0) = cP_2(0)$ folgt. Somit muss $c = 1$ gelten, falls $\theta = 0$ keine Nullstelle von P_1 bzw. P_2 ist. Da dann also für alle $\theta \in \mathbb{R}$ die Polynome $P_1(\theta)$ und $P_2(\theta)$ übereinstimmen, liefert ein Koeffizientenvergleich

$$\begin{aligned}-\sigma_{Y_1}^6\lambda_1 &= -\sigma_{Y_2}^6\lambda_2 \\ 3i\mu_{Y_1}\sigma_{Y_1}^4\lambda_1 &= 3i\mu_{Y_2}\sigma_{Y_2}^4\lambda_2 \\ (3\sigma_{Y_1}^4 + 3\mu_{Y_1}^2\sigma_{Y_1}^2)\lambda_1 &= (3\sigma_{Y_2}^4 + 3\mu_{Y_2}^2\sigma_{Y_2}^2)\lambda_2 \\ (-3i\sigma_{Y_1}^2\mu_{Y_1} - i\mu_{Y_1}^3)\lambda_1 &= (-3i\sigma_{Y_2}^2\mu_{Y_2} - i\mu_{Y_2}^3)\lambda_2,\end{aligned}$$

woraus schließlich $\sigma_{Y_1} = \sigma_{Y_2}$, $\mu_{Y_1} = \mu_{Y_2}$ und $\lambda_1 = \lambda_2$ folgt.

Wir müssen nun noch den Fall, dass $\theta = 0$ eine Nullstelle von P_1 bzw. P_2 ist, betrachten. Aus (4.25) folgt, dass $\theta_1 = 0$ für $\mu_{Y_1} = 0$ und folglich dann auch für $\mu_{Y_2} = 0$ erfüllt ist. Aus der Gleichheit der anderen Nullstellen folgt unmittelbar $\sigma_{Y_1} = \sigma_{Y_2}$ und somit aus Beziehung (4.24) auch $\lambda_1 = \lambda_2$. Die Nullstellen θ_2 und θ_3 können für Parameter aus \hat{D} nicht verschwinden, so dass wir nun alle Möglichkeiten betrachtet haben.

Wir haben gezeigt, dass für beliebige $\underline{p}_1 \in \hat{D}$ und $\underline{p}_2 \in \hat{D}$, die (4.22) erfüllen, die Beziehungen $\sigma_{Y_1} = \sigma_{Y_2}$, $\mu_{Y_1} = \mu_{Y_2}$ und $\lambda_1 = \lambda_2$ gelten. Daraus folgt dann mit (4.23), dass auch

$$i\tilde{\mu}_1\theta - \frac{\sigma_1^2}{2}\theta^2 = i\tilde{\mu}_2\theta - \frac{\sigma_2^2}{2}\theta^2 \quad \forall \theta \in \mathbb{R}$$

und schließlich $\sigma_1 = \sigma_2$ und $\tilde{\mu}_1 = \tilde{\mu}_2$ gelten. Aus $\tilde{\mu}_1 = \tilde{\mu}_2$ folgt dann mit (2.3) und $\sigma_{Y_1} = \sigma_{Y_2}$, $\mu_{Y_1} = \mu_{Y_2}$ und $\lambda_1 = \lambda_2$ auch $\mu_1 = \mu_2$ und folglich $\underline{p}_1 = \underline{p}_2$. Somit haben wir

²Wäre $\mu_{Y_1}\lambda_1 = 0$, ergäbe sich sofort $\lambda_1(\mu_{Y_1}^2 + \sigma_{Y_1}^2) = 0$, was im Widerspruch zu $\underline{p}_1 \in \hat{D}$ stände.

³Für Parameter $\underline{p} \in D \setminus \hat{D}$ würde das entsprechende Polynom P_i ($i = 1, 2$) verschwinden ($P_i(\theta) \equiv 0$). Der Fall $\sigma_{Y_1} = 0$ führt, wie wir im vorherigen Fall gesehen haben, auf eine Konstante.

gezeigt, dass für alle Parametervektoren aus \hat{D} mit $A(\underline{p}_1) = A(\underline{p}_2)$ unmittelbar $\underline{p}_1 = \underline{p}_2$ folgt, d.h. der Operator auf \hat{D} injektiv ist, womit obige Behauptung bewiesen ist \square

Folgerung 4.2.4 *Es seien $\underline{p}_1, \underline{p}_2 \in D$ zwei Parametervektoren mit $A(\underline{p}_1) = A(\underline{p}_2)$ und $\underline{p}_1 \neq \underline{p}_2$. Dann sind \underline{p}_1 und \underline{p}_2 aus $D \setminus \hat{D}$, d.h. es gilt $\lambda_1(\mu_{Y1}^2 + \sigma_{Y1}^2) = \lambda_2(\mu_{Y2}^2 + \sigma_{Y2}^2) = 0$.*

Wir zeigen nun, dass nicht beliebige verschiedene Parametervektoren aus der Menge $D \setminus \hat{D}$ gleiche Bilder erzeugen können.

Lemma 4.2.5 *Für die Parametervektoren $\underline{p}_1 = (\mu_1, \sigma_1, \lambda_1, \mu_{Y1}, \sigma_{Y1}) \in D \setminus \hat{D}$ und $\underline{p}_2 = (\mu_2, \sigma_2, \lambda_2, \mu_{Y2}, \sigma_{Y2})$ gelte $A(\underline{p}_1) = A(\underline{p}_2)$. Dann ist \underline{p}_2 ebenfalls aus der Menge $D \setminus \hat{D}$ und es gilt $\mu_1 = \mu_2$ sowie $\sigma_1 = \sigma_2$.*

Beweis. Wir haben in dem Beweis zu Theorem 4.2.3 gesehen, dass die Beziehung $A(\underline{p}_1) = A(\underline{p}_2)$ die Gleichheit der logarithmierten charakteristischen Funktionen (4.23) nach sich zieht. Wegen $\lambda_1(\mu_{Y1}^2 + \sigma_{Y1}^2) = 0$ haben wir dann

$$i\tilde{\mu}_1\theta - \frac{\sigma_1^2}{2}\theta^2 = i\tilde{\mu}_2\theta - \frac{\sigma_2^2}{2}\theta^2 + \lambda_2 \left(\exp \left(-\frac{\sigma_{Y2}^2}{2}\theta^2 + i\mu_{Y2}\theta \right) - 1 \right), \quad (4.26)$$

woraus wir unmittelbar $\lambda_2(\mu_{Y2}^2 + \sigma_{Y2}^2) = 0$ und folglich $\underline{p}_2 \in D \setminus \hat{D}$ erhalten. Koeffizientenvergleich in (4.26) liefert dann schließlich $\tilde{\mu}_1 = \tilde{\mu}_2$ und $\sigma_1 = \sigma_2$. Daraus folgt sofort $\mu_1 = \mu_2$, womit obige Aussage bewiesen ist \square

Folgerung 4.2.6 *Wenn die Operatorgleichung*

$$A(\underline{p}) = f$$

eine Lösung $\underline{p} \in \hat{D}$ besitzt, ist diese dann auch wegen Theorem 4.2.3 eindeutig. Falls $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y) \in D \setminus \hat{D}$ Lösung der Operatorgleichung ist, so folgt mit Lemma 4.2.1 und Lemma 4.2.5, dass auch $(\mu, \sigma, 0, \hat{\mu}_Y, \hat{\sigma}_Y) \in D \setminus \hat{D}$ und $(\mu, \sigma, \hat{\lambda}, 0, 0) \in D \setminus \hat{D}$ für beliebige $\hat{\mu}_Y \in \mathbb{R}$, $\hat{\sigma}_Y \geq 0$ und $\hat{\lambda} \geq 0$ Lösungen sind.

Wegen der Nichtinjektivität des Operators über dem kompletten Definitionsbereich D erfüllt die Operatorgleichung nicht die Eindeutigkeitsbedingung aus Definition 3.3.4 und ist demzufolge nicht korrekt nach Hadamard. Wir konstruieren nun ein einfaches Beispiel, so dass die Folge der Bilder $f_n = A(\underline{p}_n)$ gegen $f_0 = A(\underline{p}_0)$, die Parametervektoren \underline{p}_n aber nicht gegen \underline{p}_0 konvergieren.

Beispiel 4.2.7 *Wir betrachten die Folge $\{\underline{p}_n\}_n$ von Parametervektoren $\underline{p}_n \in D$ mit $\underline{p}_n = (\mu_n, \sigma_n, \lambda_n, \mu_{Yn}, \sigma_{Yn})$. Für die spezielle Wahl $\mu_n \equiv \mu$ und $\sigma_n \equiv \sigma$ sowie $\lambda_n = \frac{1}{n}$, $\mu_{Yn} \equiv \mu_Y > 0$ und $\sigma_{Yn} \equiv \sigma_Y > 0$ gilt $\underline{p}_n \in \hat{D}$ für jedes $n \in \mathbb{N}$. Es sei $\{f_n\}_n$ die Folge*

der Bilder der Parametervektoren, d.h. $f_n = A(\underline{p}_n)$. Ferner betrachten wir mit $\mu_0 = \mu$, $\sigma_0 = \sigma$, $\lambda_0 > 0$ und $\sigma_{Y0} = \mu_{Y0} = 0$ den Parametervektor $\underline{p}_0 = (\mu, \sigma, \lambda_0, 0, 0) \in D \setminus \hat{D}$ und dessen Bild $f_0 = A(\underline{p}_0)$.

Die Folge $\{\underline{p}_n\}_n$ konvergiert gegen $\hat{\underline{p}} = (\mu_0, \sigma_0, 0, \mu_Y, \sigma_Y) \in D \setminus \hat{D}$. Wegen der Stetigkeit des Operators A und $A(\underline{p}_0) = A(\hat{\underline{p}})$ konvergiert f_n gegen f_0 für beliebige μ_Y und σ_Y , da $\mu_n = \mu_0 = \mu$, $\sigma_n = \sigma_0 = \sigma$ und

$$\lim_{n \rightarrow \infty} \lambda_n (\mu_{Yn}^2 + \sigma_{Yn}^2) = \lim_{n \rightarrow \infty} \frac{1}{n} (\mu_Y^2 + \sigma_Y^2) = 0. \quad (4.27)$$

Für beliebiges $\lambda_0 > 0$ sowie beliebige $\mu_Y > 0$ und $\sigma_Y > 0$ haben wir aber bei Verwendung der euklidischen Norm

$$\lim_{n \rightarrow \infty} \|\underline{p}_n - \underline{p}_0\|^2 = \lim_{n \rightarrow \infty} \left(\frac{1}{n} - \lambda_0 \right)^2 + \mu_Y^2 + \sigma_Y^2 = \lambda_0^2 + \mu_Y^2 + \sigma_Y^2 > 0, \quad (4.28)$$

so dass hier die Konvergenz der Bilder f_n gegen f_0 nicht die Konvergenz der Parametervektoren \underline{p}_n gegen \underline{p}_0 nach sich zieht. Wir können den Abstand (4.28) beispielsweise mit $\mu_{Yn} = \sigma_{Yn} = n^{\frac{1}{g}}$ auch beliebig groß werden lassen, so dass Beziehung (4.27) für $g > 2$ weiterhin erfüllt ist.

Wegen der Verletzung der Injektivität für Parametervektoren aus der Menge $D \setminus \hat{D}$, können wir Folgen von Parametervektoren finden, deren Bilder gegen das Bild eines Parametervektors \underline{p}_0 aus $D \setminus \hat{D}$ konvergiert, obwohl die Folge der Parametervektoren nicht gegen \underline{p}_0 konvergiert.

Wir wollen nun die Stabilitätsbedingung der Operatorgleichung (4.20) näher untersuchen. Für Parametervektoren $\underline{p}_0 \in D \setminus \hat{D}$ ist die Eindeutigkeitsbedingung aus Definition 3.3.4 nicht erfüllt, womit der inverse Operator A^{-1} nicht existiert. Im Folgenden bezeichne

$$U(f_n) := \{\underline{p} \in D : A(\underline{p}) = f_n\}$$

die Urbildmenge der Funktion f_n . Die Werte der Sprungparameter λ_0 , μ_{Y0} und σ_{Y0} haben wegen $\lambda_0(\mu_{Y0}^2 + \sigma_{Y0}^2) = 0$ keinen Einfluss auf die Funktion $A(\underline{p}_0)$. Daher ist die Tatsache, dass wie in Beispiel 4.2.7 Folgen $\{\underline{p}_n\}$ mit $\underline{p}_n \in U(f_n)$ existieren, die nicht gegen \underline{p}_0 konvergieren (siehe (4.28)), obwohl deren Bilder f_n gegen $f_0 = A(\underline{p}_0)$ konvergieren, nicht kritisch. Die Folge $f_n \in A(D)$ konvergiere gegen $f_0 \in A(D)$ im Raum Y . Wir werden im Folgenden zeigen, dass die Diffusionsparameter μ und σ der Parametervektoren \underline{p}_n aus der Urbildmenge $U(f_n)$ unter gewissen Voraussetzungen gegen die Diffusionsparameter der Parametervektoren $\underline{p}_0 \in U(f_0)$ konvergieren, d.h.

$$\lim_{n \rightarrow \infty} \mu_n = \mu_0 \quad \text{und} \quad \lim_{n \rightarrow \infty} \sigma_n = \sigma_0. \quad (4.29)$$

Für die Sprungparameter λ , μ_Y und σ_Y können wir selbst unter diesen Zusatzvoraussetzungen nicht in jedem Fall die Konvergenz aller Parameter sichern, wir erhalten aber immer

$$\lim_{n \rightarrow \infty} \lambda_n (\mu_{Yn}^2 + \sigma_{Yn}^2) = \lambda_0 (\mu_{Y0}^2 + \sigma_{Y0}^2). \quad (4.30)$$

Falls die Sprungintensität λ gegen unendlich konvergiert, können wir allerdings keine Stabilität der Operatorgleichung (4.20) sichern, d.h. (4.29) und (4.30) müssen nicht erfüllt sein. Neben der Sprungintensität beschränken wir auch die Parameter der Sprunghöhen und betrachten für beliebige positive Konstanten λ_{\max} , $\mu_{Y_{\max}}$ und $\sigma_{Y_{\max}}$ Parametervektoren aus der Menge

$$D_{\max} := \{ \underline{p} \in D : \lambda \leq \lambda_{\max} < \infty, |\mu_Y| \leq \mu_{Y_{\max}} < \infty, \sigma_Y \leq \sigma_{Y_{\max}} < \infty \}. \quad (4.31)$$

Die Beschränkung der Sprunghöhen ist dabei keine wesentliche Einschränkung, da sowohl für $|\mu_Y| \rightarrow \infty$ als auch für $\sigma_Y \rightarrow \infty$ die Kurssprünge, und somit die absoluten Returns beliebig groß werden. Den Grenzfall $\lambda \rightarrow \infty$ werden wir noch gesondert betrachten. Um die Beziehungen (4.29) und (4.30) für Parametervektoren aus (4.31) nachzuweisen, untersuchen wir für hinreichend klein gewählte η mit $0 < \eta \leq \eta_0$ und $f_0 \in A(D_{\max})$ die unteren Niveaumengen

$$N_\eta := \{ \underline{p} \in D_{\max} : \|A(\underline{p}) - f_0\|_Y \leq \eta \}. \quad (4.32)$$

Wir zeigen zunächst, dass die unteren Niveaumengen für genügend kleine η kompakte, d.h. beschränkte und abgeschlossene Mengen im \mathbb{R}^5 sind. Dabei wollen wir uns auf die Raumwahl $Y = C(\mathbb{R})$ konzentrieren, d.h. wir messen die Abstände der Dichtefunktionen mit der Maximumnorm.

Theorem 4.2.8 *Für genügend klein gewählte η ist die Menge N_η kompakt.*

Beweis. Da wir in (4.4) die Reihenglieder der Dichtefunktion mit f_j bezeichnet haben, werden wir, um Konflikte zu vermeiden, in diesem Beweis die Folgenglieder der Reihe $\{f_n\}$ mit $f^{\{n\}}$ sowie die Grenzfunktion f_0 hier mit $f^{\{0\}}$ bezeichnen. Für Parametervektoren $\underline{p} \in D_{\max}$ gilt $0 \leq \lambda \leq \lambda_{\max} < \infty$, $|\mu_Y| \leq \mu_{Y_{\max}}$ und $0 \leq \sigma_Y \leq \sigma_{Y_{\max}} < \infty$. Im ersten Schritt zeigen wir die Abgeschlossenheit der unteren Niveaumenge N_η . Hierfür betrachten wir eine Folge $\{\underline{p}_n\}_n$ von Parametervektoren $\underline{p}_n \in N_\eta$, d.h.

$$\|A(\underline{p}_n) - f^{\{0\}}\|_{C(\mathbb{R})} \leq \eta, \quad (4.33)$$

die gegen den Parametervektor $\underline{p}_0 \in D$ konvergieren. Wegen der Stetigkeit des Operators (vgl. Theorem 4.1.3) folgt dann

$$\lim_{n \rightarrow \infty} \|A(\underline{p}_n) - A(\underline{p}_0)\|_{C(\mathbb{R})} = 0. \quad (4.34)$$

Mit (4.33) haben wir für jedes $n \in \mathbb{N}$ die Beziehung

$$\begin{aligned} \|A(\underline{p}_0) - f^{\{0\}}\|_{C(\mathbb{R})} &\leq \|A(\underline{p}_n) - A(\underline{p}_0)\|_{C(\mathbb{R})} + \|A(\underline{p}_n) - f^{\{0\}}\|_{C(\mathbb{R})} \\ &\leq \|A(\underline{p}_n) - A(\underline{p}_0)\|_{C(\mathbb{R})} + \eta, \end{aligned}$$

so dass wir schließlich mit (4.34)

$$\|A(\underline{p}_0) - f^{\{0\}}\|_{C(\mathbb{R})} \leq \eta \quad (4.35)$$

erhalten. Somit ist wegen (4.35) auch $\underline{p}_0 \in N_\eta$, womit die Abgeschlossenheit dieser Menge gezeigt ist.

Um die Beschränktheit der Menge N_η zu zeigen, müssen wir die Existenz von Schranken μ_{\max} , σ_{\min} und σ_{\max} , so dass für alle Parametervektoren $\underline{p} \in N_\eta$ die Beziehungen

$$|\mu| \leq \mu_{\max} < \infty, \quad (4.36a)$$

$$0 < \sigma_{\min} \leq \sigma \leq \sigma_{\max} < \infty, \quad (4.36b)$$

gelten, nachweisen. Wir betrachten die Parametervektoren \underline{p}_0 aus der Urbildmenge $U(f^{\{0\}}) \subset D$. Falls die Grenzfunktion $f^{\{0\}}$ Element von $A(\hat{D}) \subset A(D)$ ist, so ist die Urbildmenge wegen Theorem 4.2.3 einelementig. Andernfalls ist die Grenzfunktion $f^{\{0\}} \in A(D \setminus \hat{D}) \subset A(D)$. Wegen Lemma 4.2.5 stimmen dann die Diffusionsparameter μ_0 und σ_0 aller Parametervektoren $\underline{p}_0 = (\mu_0, \sigma_0, \lambda_0, \mu_{Y_0}, \sigma_{Y_0}) \in U(f_0)$ überein. Somit erhalten wir für alle $\underline{p}_0 \in U(f^{\{0\}})$ (vgl. (4.1))

$$f^{\{0\}}(x, \underline{p}_0) = [A(\underline{p}_0)](x) \leq \frac{1}{\sigma_0 \sqrt{2\pi\tau}} \quad (x \in \mathbb{R})$$

und

$$\begin{aligned} \|f^{\{0\}}\|_{C(\mathbb{R})} &= \max_{x \in \mathbb{R}} \left\{ \sum_{j=0}^{\infty} e^{-\lambda_0 \tau} \frac{(\lambda_0 \tau)^j}{j! \sqrt{\sigma_0^2 \tau + j \sigma_{Y_0}^2}} \phi \left(\frac{x - (\tilde{\mu}_0 \tau + j \mu_{Y_0})}{\sqrt{\sigma_0^2 \tau + j \sigma_{Y_0}^2}} \right) \right\} \\ &\geq \max_{x \in \mathbb{R}} \left\{ e^{-\lambda_0 \tau} \frac{1}{\sqrt{\sigma_0^2 \tau}} \phi \left(\frac{x - \tilde{\mu}_0 \tau}{\sqrt{\sigma_0^2 \tau}} \right) \right\} = e^{-\lambda_0 \tau} \frac{1}{\sigma_0 \sqrt{2\pi\tau}}, \end{aligned}$$

so dass

$$0 < e^{-\lambda_0 \tau} \frac{1}{\sigma_0 \sqrt{2\pi\tau}} \leq \|f^{\{0\}}\|_{C(\mathbb{R})} \leq \frac{1}{\sigma_0 \sqrt{2\pi\tau}} < \infty \quad (4.37)$$

gilt. Dabei wählen wir $\lambda_0 = 0$, falls $f^{\{0\}} \in A(D \setminus \hat{D})$, da in diesem Fall $\|f^{\{0\}}\|_{C(\mathbb{R})} = \frac{1}{\sigma_0 \sqrt{2\pi\tau}}$ gilt. Ferner können wir für alle Parametervektoren $\underline{p} \in D_{\max}$ die Norm der Funktion $A(\underline{p})$ analog zu (4.37) mit

$$e^{-\lambda_{\max} \tau} \frac{1}{\sigma \sqrt{2\pi\tau}} \leq \|A(\underline{p})\|_{C(\mathbb{R})} \leq \frac{1}{\sigma \sqrt{2\pi\tau}} \quad (4.38)$$

abschätzen. Mit (4.37) und (4.38) erhalten wir für alle Parametervektoren $\underline{p} \in N_\eta$ die Beziehung

$$e^{-\lambda_{\max} \tau} \frac{1}{\sigma \sqrt{2\pi\tau}} \leq \|A(\underline{p})\|_{C(\mathbb{R})} \leq \|A(\underline{p}) - f^{\{0\}}\|_{C(\mathbb{R})} + \|f^{\{0\}}\|_{C(\mathbb{R})} \leq \eta + \frac{1}{\sigma_0 \sqrt{2\pi\tau}},$$

und folglich

$$\sigma \geq \frac{e^{-\lambda_{\max} \tau}}{\eta \sqrt{2\pi\tau} + \frac{1}{\sigma_0}} =: \sigma_{\min} > 0. \quad (4.39)$$

Wir zeigen nun, dass σ für Parametervektoren aus der unteren Niveaumenge (4.32) für genügend kleine η nicht beliebig groß werden kann. Mit (4.37) und (4.38) gilt für Parametervektoren $\underline{p} \in N_\eta$

$$\eta \geq \|A(\underline{p}) - f^{\{0\}}\|_{C(\mathbb{R})} \geq \|f^{\{0\}}\|_{C(\mathbb{R})} - \|A(\underline{p})\|_{C(\mathbb{R})} \geq e^{-\lambda_0\tau} \frac{1}{\sigma_0\sqrt{2\pi\tau}} - \frac{1}{\sigma\sqrt{2\pi\tau}},$$

woraus

$$\sigma \leq \frac{1}{\frac{e^{-\lambda_0\tau}}{\sigma_0} - \eta\sqrt{2\pi\tau}} =: \sigma_{\max}$$

folgt. Somit können für genügend kleine η keine Parametervektoren \underline{p} mit unbeschränkten σ aus der unteren Niveaumenge (4.32) stammen. Damit haben wir (4.36b) nachgewiesen.

Wir wollen nun die Beschränktheit des Parameters μ für Parametervektoren aus N_η zeigen. Hierfür betrachten wir den Fall, dass $|\mu|$ beliebig groß wird. Für die Reihenglieder (4.4) können wir analog zum Beweis von Theorem 4.1.3 zeigen, dass für alle Parametervektoren $\underline{p} \in N_\eta$ wegen (4.31) und (4.36b) (vgl. (4.1.3)) ein endliches $\hat{j} = \hat{j}(\tilde{\varepsilon})$ existiert, so dass für alle $j \geq \hat{j}$ und ein beliebiges $\tilde{\varepsilon}$ mit $0 < \tilde{\varepsilon} < 1$ die Beziehung

$$f_j(x, \underline{p}) \leq \tilde{\varepsilon}^{j+1} \quad \forall x \in \mathbb{R}$$

gilt. Für die ersten \hat{j} Summanden der Dichtefunktion ist die Beziehung $f_j(x, \underline{p}) \leq \tilde{\varepsilon}^{j+1}$ genau dann erfüllt, wenn

$$|x - \tilde{\mu}\tau - j\mu_Y| \geq \sqrt{(\sigma^2\tau + j\sigma_Y^2)(\lambda\tau) \ln \frac{(\lambda\tau)^{2j}}{2\pi\tilde{\varepsilon}^{2(j+1)}(\sigma^2\tau + j\sigma_Y^2)^2(j!)^2}} \quad (4.40)$$

gilt. Die linke Seite von (4.40) konvergiert dann für alle $j < \hat{j}$ für $|\mu| \rightarrow \infty$ gegen unendlich, so dass für einen festen Wert x^* sofort $f_j(x^*, \underline{p}) \leq \tilde{\varepsilon}^{j+1}$ folgt. Mit $\tilde{\varepsilon} = \frac{\varepsilon}{1+\varepsilon}$ folgt dann für alle Parametervektoren aus N_η mit hinreichend großen $|\mu|$ die Beziehung

$$[A(\underline{p})](x^*) = \sum_{j=0}^{\infty} f_j(x^*, \underline{p}) \leq \sum_{j=0}^{\infty} \tilde{\varepsilon}^{j+1} = \frac{\tilde{\varepsilon}}{1-\tilde{\varepsilon}} = \varepsilon. \quad (4.41)$$

Wir wählen x^* als den Punkt, für den die Funktion $f^{\{0\}}$ ihr Maximum annimmt, was $\|f^{\{0\}}\|_{C(\mathbb{R})} = f^{\{0\}}(x^*, \underline{p}_0)$ impliziert. Da für Parametervektoren aus der Menge N_η

$$\begin{aligned} \eta &\geq \|A(\underline{p}) - f^{\{0\}}\|_{C(\mathbb{R})} = \max_{x \in \mathbb{R}} | [A(\underline{p})](x) - f^{\{0\}}(x, \underline{p}_0) | \\ &\geq | f^{\{0\}}(x^*, \underline{p}_0) - [A(\underline{p})](x^*) | \geq \|f^{\{0\}}\|_{C(\mathbb{R})} - [A(\underline{p})](x^*) \\ &\geq e^{-\lambda_0\tau} \frac{1}{\sigma_0\sqrt{2\pi\tau}} - [A(\underline{p})](x^*) \end{aligned}$$

gilt, folgt

$$[A(\underline{p})](x^*) \geq e^{-\lambda_0\tau} \frac{1}{\sigma_0\sqrt{2\pi\tau}} - \eta.$$

Somit können für genügend kleine η wegen (4.41) keine Parametervektoren \underline{p} mit unbeschränkten μ aus der unteren Niveaumenge (4.32) stammen, so dass (4.36a) gezeigt ist. Damit ist die untere Niveaumenge N_η für genügend kleine $\eta > 0$ beschränkt. Wegen der Abgeschlossenheit folgt dann, dass N_η kompakt ist, so dass obige Aussage bewiesen ist \square

Wir haben Theorem 4.2.8 für $Y = C(\mathbb{R})$ und die Maximumnorm gezeigt. Für die L^q -Räume, insbesondere $q = 2$, kann Theorem 4.2.8 mit ähnlichen Techniken gezeigt werden. Mit Hilfe der Kompaktheit der unteren Niveaumengen N_η für genügend kleine η zeigen wir nun die Stabilität der Operatorgleichung für Parametervektoren $\underline{p} \in D_{\max}$.

Theorem 4.2.9 *Es sei $\{f_n\}_n \subset A(D_{\max})$ eine Folge, die in Y gegen die Dichtefunktion $f_0 \in A(D_{\max})$ konvergiert. Dann gelten für die Parameter der Urbilder $\underline{p}_n = (\mu_n, \sigma_n, \lambda_n, \mu_{Y_n}, \sigma_{Y_n}) \in U(f_n) \cap D_{\max}$ und $\underline{p}_0 = (\mu_0, \sigma_0, \lambda_0, \mu_{Y_0}, \sigma_{Y_0}) \in U(f_0) \cap D_{\max}$ die Grenzwertbeziehungen (4.29) und (4.30). Jede unendliche Teilfolge $\{\underline{p}_{n_k}\}_k$ mit $\underline{p}_{n_k} \in U(f_{n_k}) \cap D_{\max}$ hat einen Häufungspunkt $\hat{\underline{p}} \in U(f_0) \cap D_{\max}$. Falls zusätzlich $f_0 \in A(D_{\max} \cap \hat{D})$, so sind für hinreichend große n die Urbildmengen $U(f_n) \cap D_{\max}$ sowie $U(f_0)$ einelementig und die Folge $\{\underline{p}_n\}_n$ konvergiert gegen \underline{p}_0 .*

Beweis. Wir betrachten nur Parametervektoren $\underline{p}_n \in U(f_n)$, deren Sprungintensität und Sprunghöhen beschränkt sind, d.h. $\underline{p}_n \in D_{\max} \cap U(f_n)$. Da die Folge $\{f_n\}_n$ gegen f_0 konvergiert, existiert für genügend große n ein hinreichend kleines $\eta \leq \eta_0$, so dass $\underline{p}_n \in N_\eta$ gilt und die untere Niveaumenge N_η wegen Theorem 4.2.8 kompakt ist. Wir betrachten zunächst den Fall $f_0 \in A(D_{\max} \cap \hat{D})$. Aus Theorem 4.2.3 folgt dann, dass die zugehörige Urbildmenge einelementig ist, d.h. $\{\underline{p}_0\} = U(f_0) \cap D_{\max}$ gilt. Für genügend große n müssen auch die Funktionen f_n aus $A(D_{\max} \cap \hat{D})$ stammen. Würde andernfalls eine Teilfolge $\{f_{n_k}\}_k$ mit $f_{n_k} \in A(D_{\max} \cap (D \setminus \hat{D}))$ existieren, würde für Parametervektoren $\underline{p}_{n_k} \in U(f_{n_k})$, für die

$$\lambda_{n_k} (\mu_{Y_{n_k}}^2 + \sigma_{Y_{n_k}}^2) = 0$$

gilt, wegen der Kompaktheit der Menge N_η eine Teilfolge $\{\underline{p}_{n_{k_l}}\}_l$ mit $\underline{p}_{n_{k_l}} \rightarrow \hat{\underline{p}} \in D \setminus \hat{D}$ existieren. Wegen der Stetigkeit des Vorwärtsoperators A würde dann aber $f_{n_{k_l}} = A(\underline{p}_{n_{k_l}}) \rightarrow A(\hat{\underline{p}}) \in A(D \setminus \hat{D})$ gelten, was im Widerspruch zu $f_{n_{k_l}} \rightarrow f_0$ stünde. Da also für hinreichend große n die Funktionen f_n aus der Menge $A(D_{\max} \cap \hat{D})$ stammen, sind die Urbildmengen $U(f_n) \cap D_{\max}$ ebenfalls einelementig, d.h. $\{\underline{p}_n\} = U(f_n) \cap D_{\max}$. Mit dem Satz von Tichonov über die stetige Inverse (siehe z.B. [38]) folgt dann aus der Injektivität und der Kompaktheit von N_η die Stetigkeit der inversen Abbildung $A^{-1} : A(N_\eta \cap \hat{D}) \rightarrow N_\eta \cap \hat{D}$. Somit konvergiert die Folge $\{\underline{p}_n\}_n$ gegen \underline{p}_0 , womit der zweite Teil der Aussage bewiesen ist. Aus der Konvergenz der Sprungparameter $\lim_{n \rightarrow \infty} \lambda_n = \lambda_0$, $\lim_{n \rightarrow \infty} \mu_{Y_n} = \mu_{Y_0}$ und $\lim_{n \rightarrow \infty} \sigma_{Y_n} = \sigma_{Y_0}$ folgt natürlich sofort $\lim_{n \rightarrow \infty} \lambda_n (\mu_{Y_n}^2 + \sigma_{Y_n}^2) = \lambda_0 (\mu_{Y_0}^2 + \sigma_{Y_0}^2)$.

Falls $f_0 \in A\left(D_{\max} \cap \left(D \setminus \hat{D}\right)\right)$ gilt, folgt mit Folgerung 4.2.4 und Lemma 4.2.5, dass die Diffusionsparameter aller Parametervektoren $\underline{p}_0 \in U(f_0) \cap D_{\max}$ übereinstimmen und deren Sprungparameter die Beziehung $\lambda_0(\mu_{Y_0}^2 + \sigma_{Y_0}^2) = 0$ erfüllen. Wir betrachten eine beliebige Teilfolge $\{\underline{p}_{n_k}\}_k \in N_\eta$ der Urbilder. Aus der Kompaktheit von N_η folgt, dass diese einen Häufungspunkt $\hat{\underline{p}} \in N_\eta$ besitzt. Da der Vorwärtsoperator A stetig ist, d.h. $A(\underline{p}_{n_k}) \rightarrow A(\hat{\underline{p}})$, folgt wegen $f_{n_k} \rightarrow f_0$ unmittelbar $A(\hat{\underline{p}}) = f_0$ und somit $\hat{\mu} = \mu_0$, $\hat{\sigma} = \sigma_0$ und $\hat{\lambda}(\hat{\mu}_Y^2 + \hat{\sigma}_Y^2) = 0$. Dies impliziert $\lim_{n \rightarrow \infty} \mu_n = \mu_0$, $\lim_{n \rightarrow \infty} \sigma_n = \sigma_0$ und $\lim_{n \rightarrow \infty} \lambda_n(\mu_{Y_n}^2 + \sigma_{Y_n}^2) = 0 = \lambda_0(\mu_{Y_0}^2 + \sigma_{Y_0}^2)$, womit obige Aussage bewiesen ist \square

Falls wir als Lösung nur Parametervektoren aus D_{\max} akzeptieren, so ist die Operatorgleichung (4.20) im Sinne der in Theorem 4.2.9 formulierten Eigenschaften stabil lösbar.

In Kapitel 5 werden wir anhand einer numerischen Fallstudie zeigen, dass trotz allem bei der Lösung des Least-Squares Problems (3.18) Instabilitätsprobleme auftreten. Diese Inkorrektheitsphänomene werden durch schlecht konditionierte Gleichungssysteme verursacht und können bereits bei sehr geringen Datenfehlerniveau δ auftreten.

4.3 Grenzsituationen

In Kapitel 2 haben wir bei der Einführung des Definitionsbereichs (2.4)

$$D = \{\underline{p} \in \mathbb{R}^5 : \sigma > 0, \lambda \geq 0, \sigma_Y \geq 0\}$$

den Fall $\sigma = 0$ ausgeschlossen. In diesem Fall sind die Returns

$$r_\tau = \tilde{\mu}\tau + \sum_{j=1}^{N_\tau} \ln Y_j$$

keine absolut stetigen Zufallsgrößen, so dass folglich auch keine Dichtefunktionen existieren. Im vorherigen Abschnitt haben wir gezeigt, dass die Einschränkung des Definitionsbereichs auf $D_{\max} \subset D$ die stabile Lösbarkeit der Operatorgleichung (4.20) nach sich zieht. Daher betrachten wir im Weiteren das Optimierungsproblem (3.18) nur für Parametervektoren aus $D_{\max} \subset D$, d.h.

$$\Psi(\underline{p}) = \|A(\underline{p}) - f^\delta\|_{L^2(\mathbb{R})}^2 \longrightarrow \min \quad \text{für } \underline{p} \in D_{\max}. \quad (4.42)$$

Als Folge des Ausschlusses von $\sigma = 0$ erhalten wir aber einen nicht abgeschlossenen Definitionsbereich, über dem wir in (4.42) optimieren. Wir wollen in diesem Abschnitt den Grenzfall $\sigma \rightarrow 0$ und dessen Konsequenzen für das Optimierungsproblem (4.42) untersuchen. Hierzu betrachten wir zunächst die Dichtefunktion (2.15) für den Fall, dass σ gegen null konvergiert. Für das erste Reihenglied (vgl. (4.4)) der Dichtefunktion

gilt die Grenzbeziehung

$$\begin{aligned} \lim_{\sigma \rightarrow 0} f_0(x, \underline{p}) &= \lim_{\sigma \rightarrow 0} e^{-\lambda\tau} \frac{1}{\sqrt{\sigma^2\tau}} \phi\left(\frac{x - \tilde{\mu}\tau}{\sqrt{\sigma^2\tau}}\right) \\ &= \begin{cases} \infty & \text{falls } x = \tilde{\mu}\tau \\ 0 & \text{sonst} \end{cases}, \end{aligned}$$

d.h. die Funktion f_0 strebt für $\sigma \rightarrow 0$ gegen eine δ -Distribution. Daher besitzt die Dichtefunktion mindestens ein Peak, dessen Höhe für $\sigma \rightarrow 0$ ins Unendliche wächst. Wenn wir also zu einer gegebenen beschränkten Datenfunktion f^δ das Optimierungsproblem (4.42) betrachten, sollten Parametervektoren mit sehr kleinem σ nicht als Lösung in Frage kommen.

Im vorherigen Abschnitt haben wir für Parametervektoren aus D_{\max} für die C -Norm des Operators A die Abschätzung (4.38)

$$e^{-\lambda_{\max}\tau} \frac{1}{\sigma\sqrt{2\pi\tau}} \leq \|A(\underline{p})\|_{C(\mathbb{R})} \leq \frac{1}{\sigma\sqrt{2\pi\tau}}$$

gezeigt. Sowohl die obere als auch die untere Schranke konvergieren für $\sigma \rightarrow 0$ gegen unendlich, so dass folglich auch $\|A(\underline{p})\|_{C(\mathbb{R})}$ gegen unendlich konvergiert. Analog zu (4.38) können wir

$$e^{-2\lambda_{\max}\tau} \frac{1}{2\sigma\sqrt{\tau\pi}} \leq \|A(\underline{p})\|_{L^2(\mathbb{R})} \leq \frac{1}{\sigma\sqrt{2\pi\tau}} \quad (4.43)$$

zeigen, weshalb dann für $\sigma \rightarrow 0$ auch $\|A(\underline{p})\|_{L^2(\mathbb{R})}$ gegen unendlich konvergiert. In dem Beweis zu Theorem 4.2.8 haben wir für Parametervektoren aus der unteren Niveaumenge N_η die Existenz einer oberen Schranke des Parameters σ gezeigt. Falls σ beliebig wächst, so konvergieren $\|A(\underline{p})\|_{C(\mathbb{R})}$ und $\|A(\underline{p})\|_{L^2(\mathbb{R})}$ wegen (4.38) und (4.43) gegen null. Somit können wir wegen

$$\|f^\delta\|_Y - \|A(\underline{p})\|_Y \leq \|A(\underline{p}) - f^\delta\|_Y \leq \|f^\delta\|_Y + \|A(\underline{p})\|_Y$$

neben Parametervektoren mit sehr kleinem σ auch Parametervektoren mit sehr großem σ als Lösung des Optimierungsproblems (4.42) ausschließen, da $\|A(\underline{p}) - f^\delta\|_Y$ andernfalls nicht hinreichend klein werden könnte.

Ebenso haben wir im obigen Beweis mit (4.36a) die Existenz einer Schranke μ_{\max} mit $|\mu| \leq \mu_{\max}$ gezeigt. Falls $|\mu|$ beliebig groß wird, verschiebt sich der wesentlich von null verschiedene Teil der Dichtefunktion beliebig weit nach links oder rechts auf der reellen Achse, so dass auch für solche Parametervektoren $\|f^\delta - A(\underline{p})\|_Y$ nicht beliebig klein werden kann.

Diese Überlegungen zeigen, dass die diskutierten Grenzsituationen nicht für Lösungen des Optimierungsproblems (4.42) auftreten sollten. Wir wollen dies noch etwas genauer betrachten. Die Funktion f^δ ist eine Approximation der exakten Dichtefunktion f^* , d.h. der Operator A bildet den Parametervektor $\underline{p}^* \in D_{\max}$ auf $f^* = A(\underline{p}^*)$ ab. Dabei gilt $\|f^\delta - f^*\|_Y \leq \delta$. Wenn f^δ aus den empirischen Dichtefunktionen der Returns

$r^* = (r_1^*, \dots, r_n^*)$ bzw. $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ hervorgeht, d.h. $f^\delta = h(r^*)$ bzw. $f^\delta = h(r^\delta)$, haben wir unmittelbar

$$\begin{aligned} \int_{\mathbb{R}} f^\delta(x) dx &= \int_{\mathbb{R}} \sum_q \frac{\#\{i : r_i^* \in I_q\}}{n|I_q|} \mathbb{I}_{I_q}(x) dx = \sum_q \int_{I_q} \frac{\#\{i : r_i^* \in I_q\}}{n|I_q|} dx \\ &= \sum_q \frac{\#\{i : r_i^* \in I_q\}}{n} = 1, \end{aligned}$$

sowie die Beschränktheit

$$f^\delta(x) \leq L, \quad \forall x \in \mathbb{R}$$

für eine nichtnegative Konstante L . Andernfalls fordern wir im Weiteren vergleichbare Eigenschaften.

Annahme 4.3.1 Die Datenfunktion f^δ erfülle für positive Konstanten L_1 und L_2 die folgenden Bedingungen

1. $\int_{\mathbb{R}} f^\delta(x) dx \leq L_1$,
2. $f^\delta(x) \leq L_2, \quad \forall x \in \mathbb{R}$,
3. $\|f^\delta - f^*\|_Y \leq \delta$ und
4. $f^\delta(x) \geq 0, \quad \forall x \in \mathbb{R}$.

Um Aussagen über Lösungen des Optimierungsproblems (4.42) zu treffen, betrachten wir für kleine η mit $\eta \leq \delta$ im Weiteren Parametervektoren \underline{p} aus der Menge

$$K_\eta := \left\{ \underline{p} \in D_{\max} : \|A(\underline{p}) - f^\delta\|_Y \leq \eta \right\}. \quad (4.44)$$

In Abschnitt 4.2 haben wir gezeigt, dass die vergleichbare Menge N_η für hinreichend kleine η kompakt ist. Wir können die Funktion f^δ , die die Bedingungen aus Annahme 4.3.1 erfülle, in der Menge K_η als Approximation der Funktion $f^{\{0\}}$ aus der Menge N_η interpretieren. Somit können wir mit der Argumentation aus Abschnitt 4.2 auf die Kompaktheit der Menge K_η für hinreichend kleine η schließen. Daher ist insbesondere der Diffusionsparameter σ von dem kritischen Wert null getrennt, sobald der Intensitätsparameter λ beschränkt wird. Analog zu (4.39) erhalten wir bei der Raumwahl $Y = C(\mathbb{R})$ bzw. $Y = L^2(\mathbb{R})$ für alle Parametervektoren aus K_η die Beziehung

$$\sigma \geq \frac{e^{-\lambda_{\max} \tau}}{(\delta + L_2) \sqrt{2\pi\tau}} > 0 \quad \text{bzw.} \quad \sigma \geq \frac{e^{-2\lambda_{\max} \tau}}{(\delta + L_2)^2 2\sqrt{\pi\tau}} > 0. \quad (4.45)$$

Wäre λ von oben nicht durch λ_{\max} beschränkt, würde in beiden Fällen die untere Grenze für σ gegen null konvergieren.

Wir wollen nun kurz den Einfluss einer beliebig wachsenden Sprungintensität diskutieren. Der Fall $\lambda \rightarrow \infty$ macht nur dann Sinn, falls die restlichen Sprungparameter μ_Y

und σ_Y gegen null konvergieren. Andernfalls würden die absoluten Returns $|r_\tau|$ fast sicher beliebig schnell gegen unendlich konvergieren. Wenn die Sprungintensität beliebig wächst und die absoluten Sprunghöhen beliebig klein werden, konvergiert unter gewissen Voraussetzungen der Sprungteil der Jump-Diffusions-Prozesse gegen eine geometrische Brownsche Bewegung (siehe [6]) und die Returns sind dann eine Überlagerung zweier Brownscher Bewegungen, die wiederum eine Brownsche Bewegung ist. Somit kann der Grenzfall $\sigma \rightarrow 0$ durch die gleichzeitige Konvergenz $\lambda \rightarrow \infty$ und $\mu_Y^2 + \sigma_Y^2 \rightarrow 0$ ausgeglichen werden. Ferner können für $\lambda \rightarrow \infty$ und $\mu_Y^2 + \sigma_Y^2 \rightarrow 0$ dann Parametervektoren aus der Menge $D \setminus \hat{D}$ existieren, deren Bilder übereinstimmen. Indem wir die Operatorgleichung (4.20) nur für Parametervektoren aus der Menge D_{\max} betrachten, verhindern wir das Auftreten einer solchen asymptotischen Nichtinjektivität, die auch Instabilitätsprobleme verursachen kann.

Kapitel 5

Inkorrektheitsphänomene

Die Operatorgleichung (4.20) ist für die exakte rechte Seite f^* mit $A(\underline{p}^*) = f^*$ für Parametervektoren auf dem eingeschränkten Definitionsbereich D_{\max} stabil lösbar¹. Die stabile Lösbarkeit bedeutet dabei, wie in Kapitel 4 diskutiert, dass die Konvergenz der Folge $\{f_n\}_n$ gegen f^* im Raum Y die Konvergenz der Diffusionsparameter der Lösungen \underline{p}_n gegen die Diffusionsparameter der exakten Lösung \underline{p}^* sowie der Sprungparameter im Sinne des Ausdruckes $\lambda_n (\mu_{Y_n}^2 + \sigma_{Y_n}^2)$ gegen $\lambda^* (\mu_Y^{*2} + \sigma_Y^{*2})$ nach sich zieht. Dabei haben wir die Stabilität und andere Eigenschaften des Operators A für Räume Y mit verschiedenen Normen betrachtet.

Im Weiteren werden wir uns, wie in Abschnitt 3.3 motiviert, auf den Raum $Y = L^2(\mathbb{R})$ konzentrieren. Ferner haben wir in diesem Abschnitt diskutiert, dass wir in praktischen Fällen anstatt der exakten Funktion $f^* \in A(D_{\max})$ nur die Datenfunktion f^δ mit $\|f^* - f^\delta\|_{L^2(\mathbb{R})} \leq \delta$ zur approximativen Bestimmung des Parametervektors zur Verfügung haben. Daher lösen wir anstelle der Operatorgleichung für Parametervektoren $\underline{p} \in D_{\max}$ das Least-Squares Problem (4.42), d.h. wir betrachten das Extremalproblem

$$\Psi(\underline{p}) = \|A(\underline{p}) - f^\delta\|_{L^2(\mathbb{R})}^2 \longrightarrow \min \quad \text{für } \underline{p} \in D_{\max}.$$

Wird der Datenfehler δ hinreichend klein, sollte wegen der oben diskutierten Stabilität die Lösung \underline{p}^δ eine gute Approximation der exakten Lösung \underline{p}^* darstellen.

In diesem Kapitel wollen wir anhand einiger numerischer Beispiele zeigen, dass trotz alledem bei der Lösung des Least-Squares Problems Inkorrektheitsphänomene in Form von Instabilitätsproblemen auftreten (siehe auch [19]). Im folgenden Abschnitt stellen wir eine Diskretisierung der Funktion $[A(\underline{p})](\cdot)$ zur Lösung des Least-Squares Problems (4.42) vor.

Wir suchen den Parametervektor aus der Menge D_{\max} , dessen Funktionswert Ψ minimal ist. Für dieses Optimierungsproblem benötigen wir die Kompaktheit des zulässigen Bereichs D_{\max} . Allerdings ist die Menge

$$D_{\max} = \{\underline{p} \in D : \lambda \leq \lambda_{\max} < \infty, |\mu_Y| \leq \mu_{Y_{\max}} < \infty, \sigma_Y \leq \sigma_{Y_{\max}} < \infty\}$$

¹Der Parametervektor \underline{p}^* mit $A(\underline{p}^*) = f^*$ sei dabei aus dem eingeschränkten Definitionsbereich D_{\max} .

weder beschränkt noch abgeschlossen. Im vorherigen Abschnitt haben wir aber mit Theorem 4.2.8 gezeigt, dass für genügend kleine η , $\eta \leq \delta$, die untere Niveaumenge (4.44) mit

$$K_\eta = \left\{ \underline{p} \in D_{\max} : \|A(\underline{p}) - f^\delta\|_Y \leq \eta \right\}$$

für jede Datenfunktion f^δ , welche die Bedingungen aus Annahme 4.3.1 erfüllt, eine kompakte Menge ist. Insbesondere haben wir gezeigt, dass wegen

$$\sigma \geq \frac{e^{-2\lambda_{\max}\tau}}{(\delta + L_2)^2 2\sqrt{\pi\tau}} > 0$$

(siehe rechter Ausdruck von (4.45)) immer eine untere positive Schranke für den Diffusionsparameter σ existiert. Lösen wir das Least-Squares Problem (4.42) mit einem Algorithmus, der in jeder Iteration den Funktionswert des Funktionals Ψ genügend stark verkleinert, so optimieren wir nach einer gewissen Anlaufphase automatisch über einem Kompaktum. Daher besteht keine Notwendigkeit, die Menge D_{\max} nochmals auf eine kompakte Untermenge einzuschränken. Ist andernfalls die Optimierung über einer kompakten Menge erforderlich, könnten wir wie folgt vorgehen, um das Optimierungsproblem (4.42) auf ein Kompaktum zu restringieren, ohne dabei das Minimum zu verändern.

Bemerkung 5.0.2 Wir können analog zum Beweis von Theorem 4.2.8 für ein hinreichend kleines η_0 , für das die Menge $K_{\eta_0} = \left\{ \underline{p} \in D_{\max} : \|A(\underline{p}) - f^\delta\|_{L^2(\mathbb{R})} \leq \eta_0 \right\}$ nicht leer sei, positive Konstanten μ_{\max} , σ_{\min} und σ_{\max} für die Diffusionsparameter finden (vgl. (4.36)), so dass $|\mu| \leq \mu_{\max} < \infty$ und $0 < \sigma_{\min} \leq \sigma \leq \sigma_{\max} < \infty$ für jedes $\eta \leq \eta_0$ gilt und folglich die untere Niveaumenge K_η kompakt ist. Diese Schranken sind unter anderem von δ und η_0 abhängig. Mit diesen Schranken definieren wir die kompakte Menge

$$\begin{aligned} D_{\max}^K &:= \left\{ \underline{p} \in D_{\max} : |\mu| \leq \mu_{\max}, \sigma_{\min} \leq \sigma \leq \sigma_{\max} \right\} \\ &= \left\{ \underline{p} \in \mathbb{R}^5 : -\mu_{\max} \leq \mu \leq \mu_{\max}, \sigma_{\min} \leq \sigma \leq \sigma_{\max}, \right. \\ &\quad \left. 0 \leq \lambda \leq \lambda_{\max}, -\mu_{Y_{\max}} \leq \mu_Y \leq \mu_{Y_{\max}}, 0 \leq \sigma_Y \leq \sigma_{Y_{\max}} \right\}. \end{aligned}$$

Wir bezeichnen mit \underline{p}_K^δ die Lösung des Optimierungsproblems

$$\Psi(\underline{p}) = \|A(\underline{p}) - f^\delta\|_{L^2(\mathbb{R})}^2 \longrightarrow \min \quad \text{für } \underline{p} \in D_{\max}^K.$$

Dann stimmen die Lösung \underline{p}^δ von (4.42) und \underline{p}_K^δ überein.

5.1 Diskretisiertes Problem

Wir gewinnen die Datenfunktion f^δ , wie in Abschnitt 3.1 und 3.3 beschrieben, als empirische Dichtefunktion der Returns $r^\delta = (r_1^\delta, \dots, r_n^\delta)$ bzw. $r^\star = (r_1^\star, \dots, r_n^\star)$. In praktischen Fällen, wählt man dabei häufig nicht eine Zerlegung \mathcal{I} der kompletten reellen Achse \mathbb{R} , sondern nur eines Teilintervalles $[a, b] \subset \mathbb{R}$. Die Werte der Stichprobe r^δ bzw.

r^* , die sich außerhalb des Intervalles $[a, b]$ befinden, werden dann so genannten Restklassen zugeteilt. Da die empirische Dichtefunktion dann die Form eines Histogramms hat, diskretisieren wir die Funktion $A(\underline{p}) = f(\cdot, \underline{p})$ auf diese Weise. Wir wählen die äquidistante Zerlegung $\tilde{x}_0 < \tilde{x}_1 < \dots < \tilde{x}_n$ mit $\Delta\tilde{x} = \tilde{x}_i - \tilde{x}_{i-1}$ für $i = 1, \dots, n$ und betrachten die Funktion $\tilde{f}_n : \mathbb{R} \times D \rightarrow \mathbb{R}$ mit

$$\tilde{f}_n(x, \underline{p}) = \begin{cases} 0 & x < \tilde{x}_0 \\ \frac{F(\tilde{x}_1, \underline{p})}{\Delta\tilde{x}} & \tilde{x}_0 \leq x < \tilde{x}_1 \\ \frac{F(\tilde{x}_{i+1}, \underline{p}) - F(\tilde{x}_i, \underline{p})}{\Delta\tilde{x}} & \tilde{x}_i \leq x < \tilde{x}_{i+1}, \quad i = 1, \dots, n-2 \\ \frac{1 - F(\tilde{x}_{n-1}, \underline{p})}{\Delta\tilde{x}} & \tilde{x}_{n-1} \leq x < \tilde{x}_n \\ 0 & x \geq \tilde{x}_n \end{cases} \quad (5.1)$$

als Diskretisierung der Dichtefunktion $A(\underline{p}) = f(\cdot, \underline{p})$, wobei $F(\cdot, \underline{p})$ die Verteilungsfunktion (2.14) der Returns r_τ bezeichnet. Mit (5.1) erhalten wir auf diese Weise den diskretisierten Operator \tilde{A} , der jeden Parametervektor $\underline{p} \in D$ in den zugehörigen Vektor $\underline{z} = (\tilde{f}_n(\tilde{x}_0, \underline{p}), \tilde{f}_n(\tilde{x}_1, \underline{p}), \dots, \tilde{f}_n(\tilde{x}_{n-1}, \underline{p}))^T \in \mathbb{R}^n$ abbildet. Wir diskretisieren die rechte Seite f^* der Operatorgleichung $A(\underline{p}^*) = f^*$ auf analoge Weise, und bezeichnen den Vektor mit $\underline{z}^* = (z_0^*, \dots, z_{n-1}^*)^T$ und erhalten das diskretisierte Problem

$$\tilde{A}(\underline{p}^*) = \underline{z}^*. \quad (5.2)$$

Bemerkung 5.1.1 Falls $\tilde{x}_0 \rightarrow -\infty$, $\tilde{x}_n \rightarrow \infty$ und der Durchmesser der Zerlegung $\Delta\tilde{x}$ gegen null strebt, so konvergiert die Funktionen \tilde{f}_n für $n \rightarrow \infty$ im Raum $C(\mathbb{R})$ gegen f und folglich auch im Raum $L^2(\mathbb{R})$.

Bemerkung 5.1.2 Den Übergang von dem inversen Problem (3.18) zu dem diskretisierten Problem (5.2) müssen wir mit einem Informationsverlust bezahlen. Aus der Statistik ist bekannt, dass der Kullback-Leibler Abstand zweier Wahrscheinlichkeitsmaße mit Gruppierung nie größer ist als ohne Gruppierung (siehe z.B. [52, Corollary 3.2] oder [73, S. 36]). Daher folgt für den Grenzübergang $\tilde{x}_0 \rightarrow -\infty$ und $\tilde{x}_n \rightarrow \infty$ und jedes $\Delta\tilde{x} > 0$ ein Informationsverlust durch die Diskretisierung derart, dass

$$\rho_{KL}(\tilde{f}_n(\cdot, \underline{p}_1), \tilde{f}_n(\cdot, \underline{p}_2)) \leq \rho_{KL}(f(\cdot, \underline{p}_1), f(\cdot, \underline{p}_2))$$

gilt.

Verwenden wir gestörte Daten \underline{z}^δ anstelle von \underline{z}^* , so lösen wir das diskretisierte Least-Squares Problem

$$\tilde{\Psi}(\underline{p}) = \left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2 = \sum_{i=0}^{n-1} \left(\tilde{f}_n(\tilde{x}_i, \underline{p}) - z_i^\delta \right)^2 \longrightarrow \min \quad \text{für } \underline{p} \in D_{\max}, \quad (5.3)$$

wobei wir den Abstand der Vektoren $\tilde{A}(\underline{p})$ und \underline{z}^δ mit der Euklidischen Norm $\|\cdot\|$ bewerten. Wir bezeichnen wiederum mit \underline{p}^δ die Lösung des nichtlinearen Optimierungsproblems (5.3), für dessen näherungsweise Lösung leistungsfähige numerische Optimierungsalgorithmen existieren.

5.2 Verwendete Optimierungscodes

Zur numerischen Lösung des nichtlinearen Least-Squares Problems (5.3) nutzen wir im Rahmen der durchgeführten Fallstudie die Minimierungsroutine `lsqcurvefit` der Optimierungstoolbox von Matlab Release 13. Der implementierte Algorithmus dieser Routine basiert auf einer „trust region“-Methode, bei der im $(i+1)$ -ten Iterationsschritt in einer Umgebung der Lösung der i -ten Iteration \underline{p}_i , der so genannten Vertrauenszone („trust region“), die nichtlineare Funktion $\tilde{\Psi}$ mittels Taylorentwicklung vereinfacht und für Vektoren aus der Vertrauenszone minimiert wird. In jeder Iteration muss damit ein großes lineares Gleichungssystem gelöst werden, was hier mit dem Verfahren vorkonditionierter konjugierter Gradienten (preconditioned conjugate gradient) erfolgt. Falls der Funktionswert dieser Lösung kleiner als $\tilde{\Psi}(\underline{p}_i)$ ist, d.h. einen geringeren Funktionswert als der vorherige Iterationsvektor besitzt, entspricht der neue Iterationsvektor \underline{p}_{i+1} dieser Lösung. Andernfalls wird die Vertrauensumgebung verkleinert und die linearisierte Funktion erneut bezüglich der Vektoren aus der neuen Vertrauenszone minimiert.

Die Restriktion $\underline{p} \in D_{\max}$ wird mittels einer inneren Reflexions-Newton-Methode (interior-reflective Newton method, siehe [10], [11]) berücksichtigt. Der Minimierungsroutine `lsqcurvefit` können für jeden Parameter des Parametervektors \underline{p} obere und untere Schranken übergeben werden. Falls einige Parameter nicht beschränkt werden, d.h. der Routine `lsqcurvefit` keine Schranken übermittelt werden, wählt die Routine selbst hinreichend große Schranken.

5.3 Numerische Fallstudie

Wir wollen in diesem Abschnitt die Stabilität des Optimierungsproblems (5.3) studieren. Zur Lösung dieses Least-Squares Problems müssen wir allerdings noch Grenzen λ_{\max} , $\mu_{Y_{\max}}$ sowie $\sigma_{Y_{\max}}$ für den eingeschränkten Definitionsbereich D_{\max} festlegen. In Abschnitt 4.2 genügte die Existenz dieser Schranken, so dass wir dort noch keine konkreten Werte bestimmt haben. Je kleiner wir diese Grenzen wählen², desto besser werden die numerischen Ergebnisse, so dass die Wahl dieser Schranken die Güte der Ergebnisse beeinflusst. Da das Ziel dieser Fallstudie die Analyse des Stabilitätsverhaltens des diskretisierten inversen Problems ist, betrachten wir hier den für die Stabilität schlechtesten Fall $\lambda_{\max} = \mu_{Y_{\max}} = \sigma_{Y_{\max}} = \infty$, d.h. wir wählen für diese Fallstudie $D_{\max} = D$ und lösen die diskrete Variante des ursprünglichen Least-Squares Problems (3.18), für das wir alle Parametervektoren $\underline{p} \in D$ als Lösung akzeptieren.

Normalerweise bestimmen wir die Datenfunktion f^δ bzw. den Vektor \underline{z}^δ aus den empirischen Returns, so dass eine sinnvolle Einschätzung des Datenfehlerniveaus δ schwierig ist. Deshalb werden wir zur Analyse der Stabilität den Vektor $\underline{z}^* = \tilde{A}(p^*)$ komponentenweise mit normalverteilten und unkorrelierten Fehlern $\epsilon_i \sim N(0, \delta^2)$ mittels $z_i^\delta = z_i^*(1 + \epsilon_i)$ stören, so dass $\mathbb{E} \|\underline{z}^* - \underline{z}^\delta\|^2 = \sum_{i=0}^{n-1} (z_i^*)^2 \mathbb{E}(\epsilon_i^2) = \delta^2 \|\underline{z}^*\|^2$ und folglich

²Würden wir einen sehr kleinen Wert $\mu_{Y_{\max}}$ mit $\underline{p}^* \in D_{\max}$ wählen, so wäre der Fehler des geschätzten Parameters μ_Y^δ unabhängig vom Datenfehler δ wegen $|\mu_Y^\delta - \mu_Y^*| \leq 2\mu_{Y_{\max}}$ ebenfalls klein.

$\mathbb{E} \left(\frac{\|\underline{z}^* - \underline{z}^\delta\|^2}{\|\underline{z}^*\|^2} \right) = \delta^2$ gelten. Den Vektor \underline{z}^δ können wir als Skelett der Datenfunktion f^δ interpretieren. Abbildung 5.1 illustriert die verwendeten Funktionen und Daten.

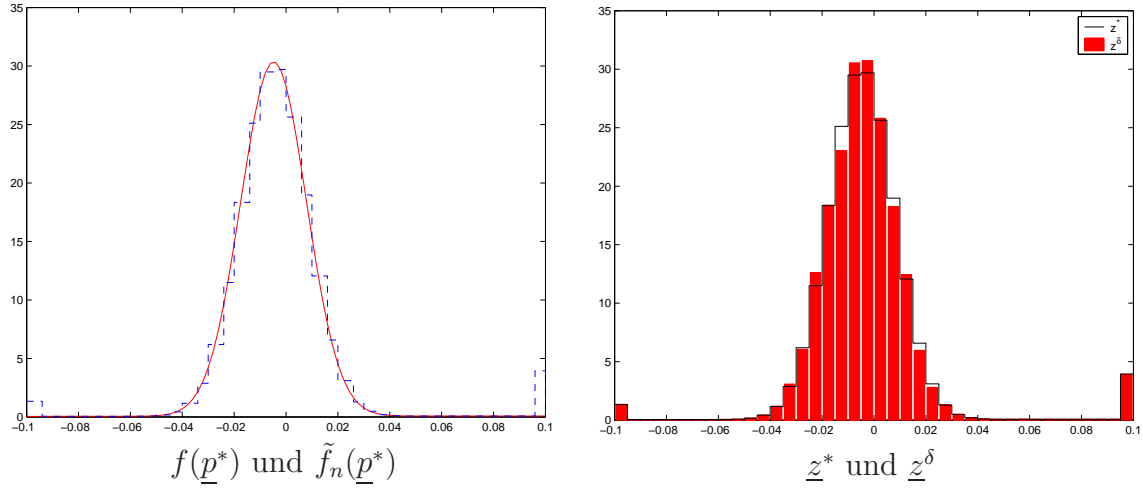


Abbildung 5.1: Diskretisierte und gestörte Daten

Wir wollen zur Einschätzung der Lösung \underline{p}^δ von (5.3) die zugehörigen Semiinvarianten $s_{\tau,k}(\underline{p}^\delta)$, die wir mit Formel (2.19) berechnen, mit den empirischen Semiinvarianten der gestörten Daten $s_{\tau,k}^\delta$ und den exakten Semiinvarianten $s_{\tau,k}(\underline{p}^*)$ vergleichen. Die Momente und Semiinvarianten enthalten die gleichen Informationen und können wegen Satz 2.2.8 ineinander umgerechnet werden. Wegen Lemma 2.2.10 sind die Semiinvarianten bezüglich der Zeit skalierbar, weshalb wir diese bevorzugt betrachten. In Abschnitt 2.2.2 haben wir gezeigt, wie wir mittels (2.21) die empirischen Semiinvarianten $s_{\tau,k}^\delta$ aus den empirischen zentralen Momenten berechnen können. Üblicherweise bestimmen wir den Vektor \underline{z}^δ aus den empirischen Returns r^δ . Für diese können wir dann auch mittels statistischer Schätzverfahren die empirischen Momente und somit die empirischen Semiinvarianten berechnen. Da wir hier den Datenvektor \underline{z}^δ nicht aus den empirischen Returns, sondern wie oben beschrieben zufällig generieren, müssen wir die empirischen Semiinvarianten auf einen anderen Wege berechnen.

Falls wir das Intervall $[\tilde{x}_0, \tilde{x}_n]$ genügend groß und den Durchmesser $\Delta\tilde{x}$ der Zerlegung hinreichend klein wählen, können wir die empirischen Momente $m_{\tau,k}^\delta$ mittels

$$m_{\tau,k}(\underline{p}) = \int_{\mathbb{R}} x^k f(x, \underline{p}) dx \approx \int_{\mathbb{R}} x^k \tilde{f}_n(x, \underline{p}) dx = \sum_{i=0}^{n-1} z_i \frac{1}{k+1} (\tilde{x}_{i+1}^{k+1} - \tilde{x}_i^{k+1})$$

näherungsweise berechnen. Allerdings ist diese Wahl der Parameter \tilde{x}_0 , \tilde{x}_n und $\Delta\tilde{x}$ ungünstig zur Lösung des Least-Squares Problems (5.3), da sie die Lösung hochdimensionaler nichtlinearer Gleichungssysteme erfordert. Außerdem werden die Funktionswerte der Dichtefunktion in den Schwänzen für jeden beliebigen Parametervektor $\underline{p} \in D_{\max}$ sehr klein, weshalb sich die Stabilität des Gleichungssystem verschlechtert. Falls \tilde{x}_0 gegen minus unendlich und \tilde{x}_n gegen unendlich konvergiert, so konvergiert $\mathbb{P}(r_\tau \in [\tilde{x}_0, \tilde{x}_n])$ gegen eins, d.h. für sehr große Absolutwerte \tilde{x}_0 und \tilde{x}_n deckt das

Intervall $[\tilde{x}_0, \tilde{x}_n]$ mit nahezu 100-prozentiger Wahrscheinlichkeit jede Realisierung der Zufallsgröße r_τ ab. Wählen wir \tilde{x}_0 und \tilde{x}_n nicht beliebig klein bzw. groß, so ist diese Wahrscheinlichkeit kleiner als eins. Allerdings ist die Funktion \tilde{f}_n so konstruiert, dass die Funktionswerte des Intervalls $[\tilde{x}_0, \tilde{x}_1]$ die Wahrscheinlichkeitsmasse des Intervalls $(-\infty, \tilde{x}_1)$ und die Funktionswerte des Intervalls $[\tilde{x}_{n-1}, \tilde{x}_n]$ die Wahrscheinlichkeitsmasse des Intervalls $[\tilde{x}_{n-1}, \infty)$ enthalten, so dass es unproblematisch ist, wenn das Intervall $[\tilde{x}_0, \tilde{x}_n]$ nicht die komplette reelle Achse überdeckt. Allerdings stimmen dann die Funktionen \tilde{f}_n und f auf den Intervallen $[\tilde{x}_0, \tilde{x}_1]$ und $[\tilde{x}_{n-1}, \tilde{x}_n]$ nicht gut überein³, so dass dann auch obige Approximation keine gute Näherung der Momente darstellt. Um die empirischen Momente bzw. Semiinvarianten direkt aus der Datenfunktion f^δ bzw. dem zugehörigen Datenvektor \underline{z}^δ besser zu berechnen⁴, führen wir eine zweite Zerlegung $\hat{x}_0, \dots, \hat{x}_m$ mit $\Delta\hat{x} = \Delta\tilde{x}$ ein, die die Zerlegung $\tilde{x}_0, \dots, \tilde{x}_n$ erweitert, d.h. für $m > n$ gilt $[\tilde{x}_0, \tilde{x}_n] \subsetneq [\hat{x}_0, \hat{x}_m]$ und alle Punkte $\hat{x}_i \in [\tilde{x}_0, \tilde{x}_n]$ stimmen mit dem entsprechenden \tilde{x}_j überein. Wir führen die Funktion $\hat{f}_m : \mathbb{R} \times D \rightarrow \mathbb{R}$ mit

$$\hat{f}_m(x, \underline{p}) = \begin{cases} 0, & x < \hat{x}_0 \\ \frac{F(\hat{x}_{j+1}, \underline{p}) - F(\hat{x}_j, \underline{p})}{\Delta\hat{x}}, & \hat{x}_i \leq x < \hat{x}_{j+1} \quad j = 0, \dots, m-1 \\ 0, & x \geq \hat{x}_m \end{cases}$$

ein, die für alle $x \in [\tilde{x}_1, \tilde{x}_{n-1})$ mit \tilde{f}_n übereinstimmt. Wie in Abbildung 5.2 deutlich zu

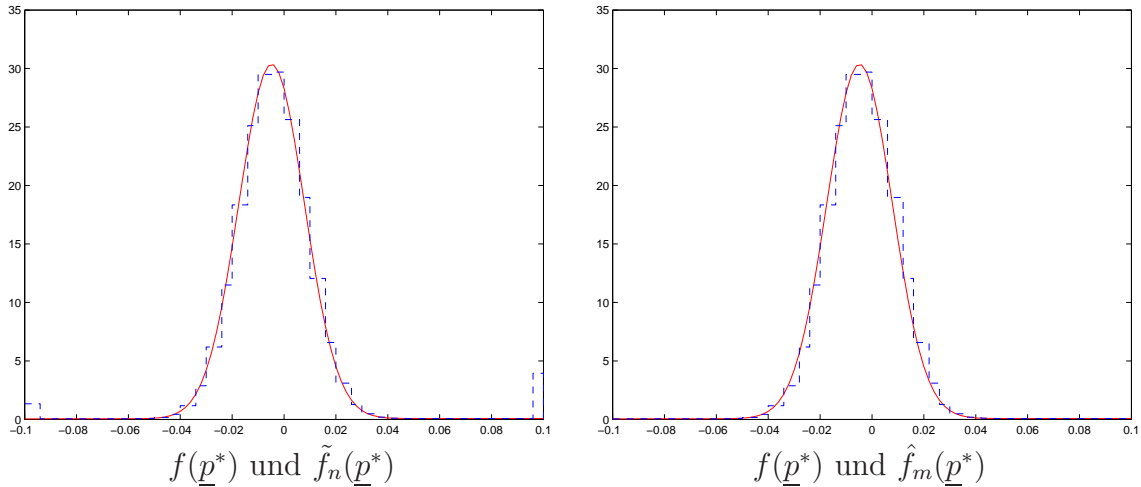


Abbildung 5.2: Diskretisierung zur Lösung des Least-Squares Problems und zur Berechnung der empirischen Momente

erkennen ist, unterscheiden sich die beiden Funktionen hauptsächlich in den Randintervallen $[\tilde{x}_0, \tilde{x}_1)$ und $[\tilde{x}_{n-1}, \tilde{x}_n)$ der inneren Zerlegung $\tilde{x}_0, \dots, \tilde{x}_n$. Während die Funktion \tilde{f}_n die Restwahrscheinlichkeit der Menge $\mathbb{R} \setminus [\tilde{x}_0, \tilde{x}_n]$ in den Intervallen $[\tilde{x}_0, \tilde{x}_1)$ bzw.

³Je kleiner \tilde{x}_0 und je größer \tilde{x}_n gewählt sind, desto kleiner werden die Restwahrscheinlichkeiten, die in den Funktionswerten von \tilde{f}_n auf den Intervallen $[\tilde{x}_0, \tilde{x}_1)$ und $[\tilde{x}_{n-1}, \tilde{x}_n)$ enthalten sind, so dass dann auch die Abweichungen zwischen \tilde{f}_n und f immer kleiner werden.

⁴Wir benötigen diesen Schritt nur für diese Fallstudie, da wir bei praktischen Problemen die empirischen Momente bzw. Semiinvarianten direkt aus den empirischen Returns berechnen.

$[\tilde{x}_{n-1}, \tilde{x}_n)$ enthält, wird die Restwahrscheinlichkeit der Menge $\mathbb{R} \setminus [\hat{x}_0, \hat{x}_m]$, die sehr klein ist, bei der Funktion \hat{f}_m nicht berücksichtigt.

Mit der Funktion \hat{f}_m und dem analog zu $\underline{z}^\delta \in \mathbb{R}^n$ gestörten Vektor $\underline{\hat{z}}^\delta \in \mathbb{R}^m$ des exakten Vektors $\underline{\hat{z}} = (\hat{f}_m(\hat{x}_0, \underline{p}^*), \hat{f}_m(\hat{x}_1, \underline{p}^*), \dots, \hat{f}_m(\hat{x}_{m-1}, \underline{p}^*))$ berechnen wir dann mittels

$$m_{\tau,k}^\delta := \sum_{j=0}^{m-1} \frac{1}{k+1} (\hat{x}_{j+1}^{k+1} - \hat{x}_j^{k+1}) \hat{z}_j^\delta \quad (5.4)$$

die empirischen Momente.

Da die Funktionen \tilde{f}_n und \hat{f}_m für alle x aus $[\tilde{x}_1, \tilde{x}_{n-1}]$ übereinstimmen, sind auch die entsprechenden Komponenten der Vektoren $\underline{\hat{z}}^\delta$ und \underline{z}^δ identisch. Die aus dem Vektor $\underline{\hat{z}}^\delta$ berechneten empirischen Momente (5.4) nutzen wir dann, um mittels (2.20) die gewünschten Semiinvarianten $s_{\tau,k}^\delta$ zu berechnen.

Bemerkung 5.3.1 Die Zerlegung $\hat{x}_0, \dots, \hat{x}_m$ ist eine Erweiterung der Zerlegung $\tilde{x}_0, \dots, \tilde{x}_n$, so dass wir zur Berechnung der empirischen Momente in (5.4) nur noch die Werte \hat{x}_0 und \hat{x}_m bestimmen müssen. Das Intervall $[\hat{x}_0, \hat{x}_m]$ sollte möglichst so groß sein, dass $\mathbb{P}(r_\tau \in [\hat{x}_0, \hat{x}_m]) \approx 1$ und folglich (5.4) eine gute Approximation der exakten Momente darstellt. Daher wählen wir die Intervallgrenzen so, dass $\mathbb{E}r_\tau - 5\sqrt{\mathbb{D}^2 r_\tau} \geq \hat{x}_0$ und $\mathbb{E}r_\tau + 5\sqrt{\mathbb{D}^2 r_\tau} \leq \hat{x}_m$ erfüllt sind.

Wir werden nun anhand eines numerischen Beispiels auftretende Inkorrekttheitsphänomene zeigen, indem wir das Least-Squares Problem (5.3) sowohl für exakte als auch für leicht gestörte Daten numerisch lösen. Die ungestörten Daten erhalten wir, indem wir den Datenfehler $\delta_1 = 0$ wählen. Mit einem kleinen $\delta_2 > 0$ bekommen wir leicht gestörte Daten.

5.3.1 Verwendete Daten

Der für die erste Fallstudie verwendete Parametervektor \underline{p}^* sowie die weiteren Einstellungen sind in Tabelle 5.1 dargestellt.

μ	σ	λ	μ_Y	σ_Y	τ	δ_1	δ_2	n	\tilde{x}_0	\tilde{x}_n
0.1	0.2	10	0.1	0.2	0.004	0	0.01	40	-0.1	0.1

Tabelle 5.1: Werte des Parametervektors \underline{p}^* und weitere Einstellungen

Bemerkung 5.3.2 Wir haben in dieser Fallstudie $\tau = \frac{1}{250} = 0.004$ gewählt, d.h. wir simulieren Tagesreturns bzw. deren empirische Dichtefunktion. In praxi erhalten wir diesen Zeitabstand beispielsweise bei der Analyse von Tagesschlusskursen.

Eine zu diesen Werten simulierte Preistrajektorie S ist in der linken Graphik von Abbildung 5.3 zu sehen. Die rechte Seite zeigt die Daten \underline{z}^* und \underline{z}^{δ_2} bzw. die zugehörigen stückweise konstanten Funktionen \tilde{f}_n und $\tilde{f}_n^{\delta_2}$. Letztere Funktion haben wir als Histogramm dargestellt.

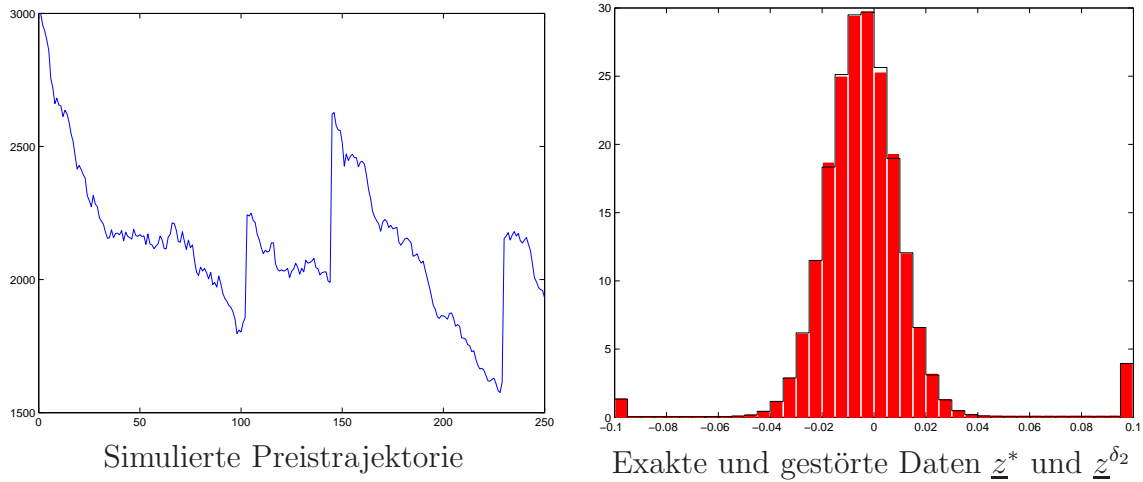


Abbildung 5.3: Simulierter Kursverlauf und verwendete Daten

Die empirischen Momente $m_{\tau,k}^{\delta_1}$, die wir mittels (5.4) berechnen, und die Semiinvarianten $s_{\tau,k}^{\delta_1}$ der ungestörten Daten zeigen kaum Abweichungen zu den exakten Werten $m_{\tau,k}(\underline{p}^*)$ bzw. $s_{\tau,k}(\underline{p}^*)$. Deswegen haben wir diese nicht nochmals abgebildet.

k	Semiinvarianten			Momente		
	$s_{\tau,k}(\underline{p})$	$s_{\tau,k}^{\delta_2}$	Abw.	$m_{\tau,k}(\underline{p})$	$m_{\tau,k}^{\delta_2}$	Abw.
1	-0.000779874	-0.000766772	1.68%	-0.000779874	-0.000766772	1.68%
2	+0.002160000	+0.002162937	0.14%	+0.002160608	+0.002163525	0.13%
3	+0.000520000	+0.000519441	0.11%	+0.000514946	+0.000514465	0.09%
4	+0.000292000	+0.000291870	0.04%	+0.000304383	+0.000304319	0.02%
5	+0.000112400	+0.000112369	0.03%	+0.000122442	+0.000122434	0.01%
6	+0.000069640	+0.000069666	0.04%	+0.000081380	+0.000081419	0.05%
7	+0.000033940	+0.000033959	0.06%	+0.000044162	+0.000044186	0.06%
8	+0.000022893	+0.000022913	0.09%	+0.000033537	+0.000033562	0.07%

Tabelle 5.2: Empirische und exakte Semiinvarianten sowie Momente

Die Abweichungen zwischen $m_{\tau,k}^{\delta_2}$ und $m_{\tau,k}(\underline{p}^*)$ sowie zwischen $s_{\tau,k}^{\delta_2}$ und $s_{\tau,k}(\underline{p}^*)$ liegen im gleichen Bereich wie der Datenfehler δ_2 , wie in Tabelle 5.2 zu sehen ist.

5.3.2 Ergebnisse im ungestörten Fall

Der in Abschnitt 5.2 beschriebene iterative Optimierungsalgorithmus zur Lösung des Least-Squares Problems liefert im ungestörten Fall, d.h. $\delta_1 = 0$, mit der Startlösung $\underline{p}_0 = (0, 0.001, 0, 0, 1)^T$ sehr gute Ergebnisse, die wir in Tabelle 5.3 dargestellt haben.

Der geschätzte Parametervektor \underline{p}^{δ_1} ist eine sehr gute Approximation des exakten Parametervektors \underline{p}^* . Daher stimmen die diskretisierten Dichtefunktionen $\tilde{f}_n(\cdot, \underline{p}^{\delta_1})$ und $\tilde{f}_n(\cdot, \underline{p}^*)$ sehr gut überein, und die Abweichungen zwischen den Momenten bzw. Semiinvarianten, die in Tabelle 5.4 abgebildet sind, fallen sehr gering aus. Wir stellen die

Parameter	\underline{p}^{δ_1} (geschätzt)	\underline{p}^* (exakt)	\underline{p}_0 (Startwert)	Abweichung
μ	0.1022	0.1	0.0	2.24%
σ	0.2000	0.2	0.001	0.00%
λ	9.9920	10.0	0.0	0.08%
μ_Y	0.1002	0.1	0.0	0.22%
σ_Y	0.2003	0.2	1.0	0.17%
$\ \tilde{f}_n(\underline{p}^*) - \tilde{f}_n(\underline{p}^{\delta_1})\ _{L^2(\mathbb{R})} = 0.000096$				

Tabelle 5.3: Geschätzte Parameter im ungestörten Fall

Momente und Semiinvarianten der Übersichtlichkeit halber für jährliche Returns, d.h. $\tau = 1$, dar.

k	$s_{1,k}(\underline{p}^{\delta_1})$	$s_{1,k}(\underline{p}^*)$	Abw.	$m_{1,k}(\underline{p}^{\delta_1})$	$m_{1,k}(\underline{p}^*)$	Abw.
1	-0.1935712	-0.1949685	0.72%	-0.1935712	-0.1949685	0.72%
2	+0.5414362	+0.5400000	0.27%	+0.5789060	+0.5780127	0.15%
3	+0.1306442	+0.1300000	0.50%	-0.1910282	-0.1932603	1.15%
4	+0.0734753	+0.0730000	0.65%	+0.9749081	+0.9710226	0.40%
5	+0.0283397	+0.0281000	0.85%	-0.1771991	-0.1847419	4.08%
6	+0.0175866	+0.0174100	1.01%	+2.8394970	+2.8185076	0.74%
7	+0.0085878	+0.0084850	1.21%	+0.7387303	+0.6950009	6.29%
8	+0.0058021	+0.0057233	1.38%	+12.6582942	+12.5104716	1.18%

Tabelle 5.4: Semiinvarianten und Momente des geschätzten Parametervektors \underline{p}^{δ_1}

Die guten Ergebnisse im ungestörten Fall zeigen, dass der von uns gewählte Zugang zur Lösung des inversen Problems vernünftig gewählt ist. Außerdem können wir davon ausgehen, dass die Optimierungsroutine zur Lösung von (5.3) mit den Einstellungen aus Tabelle 5.1 sowie dem Startvektor \underline{p}_0 zuverlässig funktioniert.

Wählen wir eine feinere Diskretisierung, d.h. erhöhen wir n , so werden die Abweichungen zwischen dem exakten Parametervektor \underline{p}^* und dem geschätztem Parametervektor \underline{p}^{δ_1} noch geringer. Allerdings kostet die feinere Diskretisierung deutlich mehr Rechenzeit. Falls wir zusätzlich zum Diskretisierungsgrad n auch das Intervall $[\tilde{x}_0, \tilde{x}_n]$ vergrößern, tendieren die Abweichungen gegen null, wobei jedoch die Rechenzeit des Programms sich nochmals deutlich erhöht. Wir werden im Weiteren die Einstellungen aus Tabelle 5.3 verwenden, da obiger Optimierungsalgorithmus dort bereits sehr gute Ergebnisse in einer akzeptablen Rechenzeit liefert.

In Kapitel 4 haben wir gezeigt, dass der Vorwärtsoperator A eine gewisse glättende Wirkung hat, so dass verschiedene Parametervektoren auf ähnliche Dichtefunktionen abgebildet werden können. Daher variieren die Lösungen \underline{p}^{δ_1} stark in Abhängigkeit von \underline{p}_0 . So liefert beispielsweise der Startvektor $\underline{p}_0 = (1, 1, 1, 1, 1)$ keine befriedigende Approximation \underline{p}^{δ_1} des exakten Parametervektors \underline{p}^* , obwohl die diskretisierten Dichtefunktion $\tilde{f}_n(\cdot, \underline{p}^{\delta_1})$ die Funktion $\tilde{f}_n(\cdot, \underline{p}^*)$ sehr gut approximiert. Der Übersichtlichkeit halber verzichten wir hier auf eine Illustration dieser Abhängigkeit und wählen für den leicht gestörten Fall einen Startvektor, der im ungestörten Fall sehr gute Ergebnisse

geliefert hat. Wir wollen nun die Ergebnisse diskutieren, die wir bei der Verwendung des leicht gestörten Datenvektors \underline{z}^{δ_2} erhalten.

5.3.3 Ergebnisse im leicht gestörten Fall

Obwohl der Datenfehler nur ein Prozent beträgt, tritt eine relativ große Abweichung zwischen dem exakten Parametervektor \underline{p}^* und dem geschätzten Parametervektor \underline{p}^{δ_2} auf. Dabei haben wir das Least-Squares Problem (5.3) mit dem gleichen Optimierungsalgorithmus sowie der gleichen Startlösung \underline{p}_0 (siehe Tabelle 5.5) wie im ungestörten Fall gelöst.

Parameter	\underline{p}^{δ_1} (geschätzt)	\underline{p}^* (exakt)	\underline{p}_0 (Startwert)	Abweichung
μ	-0.0424	+0.1	0.0	142.35%
σ	+0.2003	+0.2	0.001	0.16%
λ	+10.6748	+10.0	0.0	6.75%
μ_Y	+0.0865	+0.1	0.0	13.54%
σ_Y	+0.1784	+0.2	1.0	10.82%
$\ \tilde{f}_n(\underline{p}^*) - \tilde{f}_n(\underline{p}^{\delta_2})\ _{L^2(\mathbb{R})} = 0.030126$				

Tabelle 5.5: Geschätzte Parameter im gestörten Fall

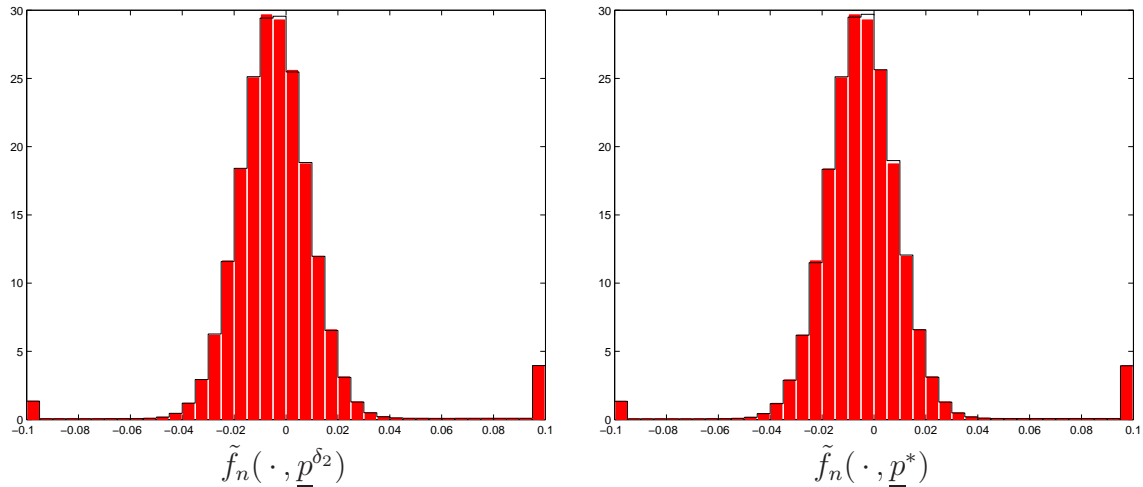


Abbildung 5.4: Histogramm der gestörten Daten y^{δ_2} (rot) und diskretisierte Dichtefunktionen des exakten sowie des geschätzten Parametervektors im Vergleich

Vor allem der Diffusionsparameter μ reagiert sehr sensibel auf leichte Datenschwankungen. Dennoch approximiert, wie in Abbildung 5.4 zu sehen ist, die diskretisierte Dichtefunktion des geschätzten Parametervektors $\tilde{f}_n(\cdot, \underline{p}^{\delta_2})$ die diskretisierte Dichtefunktion des exakten Parametervektors $\tilde{f}_n(\cdot, \underline{p}^*)$ sehr gut. In Tabelle 5.6 sind die Semiinvarianten und Momente des geschätzten Parametervektors sowie zum Vergleich die Werte des exakten Parametervektors abgebildet. Interessanterweise treten hier recht deutliche Abweichungen zwischen den exakten und geschätzten Werten auf. Diesen Effekt werden wir später zur Regularisierung des diskretisierten inversen Problems (5.2) verwenden.

k	$s_{1,k}(\underline{p}^{\delta_2})$	$s_{1,k}(\underline{p}^*)$	Abw.	$m_{1,k}(\underline{p}^{\delta_2})$	$m_{1,k}(\underline{p}^*)$	Abw.
1	-0.2900873	-0.1949685	48.79%	-0.2900873	-0.1949685	48.79%
2	+0.4594911	+0.5400000	14.91%	+0.5436417	+0.5780127	5.95%
3	+0.0949759	+0.1300000	26.94%	-0.3293127	-0.1932603	70.40%
4	+0.0482313	+0.0730000	33.93%	+0.8105025	+0.9710226	16.53%
5	+0.0162548	+0.0281000	42.15%	-0.5702948	-0.1847419	208.70%
6	+0.0090766	+0.0174100	47.87%	+1.9624666	+2.8185076	30.37%
7	+0.0038871	+0.0084850	54.19%	-1.1839089	+0.6950009	270.35%
8	+0.0023572	+0.0057233	58.81%	+6.7970793	+12.5104716	45.67%

Tabelle 5.6: Semiinvarianten und Momente des geschätzten Parametervektors \underline{p}^{δ_2}

	μ	σ	λ	μ_Y	σ_Y	L^2 -Abstand
\underline{p}^*	0.1	0.2	10.0	0.1	0.2	0.0
\underline{p}_0	0.0	0.001	0.0	0.0	1.0	
Sim 1	0.6413	0.1993	8.6430	0.1522	0.2792	0.028726
Abw.	5.4e+02%	3.4e-01%	1.4e+01%	5.2e+01%	4.0e+01%	
Sim 2	-0.1225	0.1992	11.4048	0.0759	0.1617	0.039022
Abw.	2.2e+02%	4.1e-01%	1.4e+01%	2.4e+01%	1.9e+01%	
Sim 3	-0.0913	0.2003	11.1035	0.0803	0.1685	0.043212
Abw.	1.9e+02%	1.4e-01%	1.1e+01%	2.0e+01%	1.6e+01%	
Sim 4	0.0354	0.2003	10.3857	0.0924	0.1876	0.031076
Abw.	6.5e+01%	1.5e-01%	3.9e+00%	7.6e+00%	6.2e+00%	
Sim 5	0.2557	0.199	9.5495	0.1128	0.2198	0.029238
Abw.	1.6e+02%	3.7e-02%	4.5e+00%	1.3e+01%	9.9e+00%	
Sim 6	0.5253	0.1995	8.8196	0.1415	0.2637	0.037103
Abw.	4.3e+02%	2.4e-01%	1.2e+01%	4.1e+01%	3.2e+01%	

Tabelle 5.7: Ergebnisse für verschiedene Simulationen der gestörten Daten

Wir haben noch sechs weitere gestörte Datenvektoren \underline{z}^{δ_2} simuliert und für diese das Optimierungsproblem (5.3) gelöst. Die in Tabelle 5.7 dargestellten Ergebnisse der geschätzten Parameter sowie der Abweichungen $\left\| \tilde{f}_n(\underline{p}^*) - \tilde{f}_n(\underline{p}^{\delta_2}) \right\|_{L^2(\mathbb{R})}$ zeigen die Instabilität des diskretisierten inversen Problems, die bereits bei einem Datenfehlerniveau von nur einem Prozent auftreten. Daher benötigen wir einen Regularisierungszugang, um diese numerische Instabilität zu überwinden und adäquate Approximationen \underline{p}^{δ_2} von \underline{p}^* zu erhalten.

Zu Beginn dieses Kapitels haben wir die Verwendung der oberen Grenzen $\lambda_{\max} = \infty$, $\mu_{Y_{\max}} = \infty$ sowie $\sigma_{Y_{\max}} = \infty$ für diese Fallstudie motiviert, so dass hier der eingeschränkte Definitionsbereich D_{\max} mit dem üblichen Definitionsbereich D übereinstimmt. Da bei keinem der oben aufgeführten Beispiele die geschätzten Parameterwerte ungewöhnlich große Werte angenommen haben, scheint eine Beschränkung der Sprungparameter nicht notwendig zu sein. Bei Rechnungen mit verschiedenen endlichen Schranken λ_{\max} , $\mu_{Y_{\max}}$ und $\sigma_{Y_{\max}}$ haben die approximierten Lösungen leider keine wesentlich bessere Stabilität aufgewiesen, so dass eine vernünftige und nicht zu enge Ein-

schränkung des Definitionsbereichs D auf D_{\max} zumindest praktisch keine Stabilität impliziert. Um diese Fallstudie nicht unnötig auszudehnen, verzichten wir an dieser Stelle auf eine detaillierte Darstellung dieser Ergebnisse.

Bisher haben wir immer das diskretisierte Least-Squares Problem für $n = 40$ Klassen betrachtet, was einer Diskretisierungslänge von $\Delta\tilde{x} = 0.005$ entspricht. Eine feinere Diskretisierung liefert eine bessere Approximation der Dichtefunktion und somit sollte \underline{p}^{δ_2} auch eine bessere Approximation von \underline{p}^* darstellen. Eine Erhöhung der Klassen auf $n = 100$ bzw. $n = 200$ liefert mit $\Delta\tilde{x} = 0.002$ bzw. $\Delta\tilde{x} = 0.001$ eine deutlich feinere Diskretisierung.

Parameter	\underline{p}^{δ_2} (geschätzt)	\underline{p}^* (exakt)	\underline{p}_0 (Startwert)	Abweichung
μ	+0.3846	+0.1	0.0	284.56%
σ	+0.1993	+0.2	0.001	0.36%
λ	+9.1030	+10.0	0.0	8.97%
μ_Y	+0.1285	+0.1	0.0	28.46%
σ_Y	+0.2444	+0.2	1.0	22.19%
$\ \tilde{f}_n(\underline{p}^*) - \tilde{f}_n(\underline{p}^{\delta_2})\ _{L^2(\mathbb{R})} = 0.02133$				

Tabelle 5.8: Geschätzte Parameter für $n = 100$ Klassen

Parameter	\underline{p}^{δ_2} (geschätzt)	\underline{p}^* (exakt)	\underline{p}_0 (Startwert)	Abweichung
μ	+12.625	+0.1	0.0	12525.1%
σ	+0.2016	+0.2	0.001	0.78%
λ	+7.1072	+10.0	0.0	28.93%
μ_Y	+0.6034	+0.1	0.0	503.4%
σ_Y	+0.9752	+0.2	1.0	387.6%
$\ \tilde{f}_n(\underline{p}^*) - \tilde{f}_n(\underline{p}^{\delta_2})\ _{L^2(\mathbb{R})} = 0.035186$				

Tabelle 5.9: Geschätzte Parameter für $n = 200$ Klassen

Tabelle 5.8 und Tabelle 5.9 zeigen die entsprechenden Ergebnisse, die wir für jeweils eine Simulation des zufällig gestörten Datenvektors $\underline{z}^{\delta_2} \in \mathbb{R}^{100}$ bzw. $\underline{z}^{\delta_2} \in \mathbb{R}^{200}$ berechnet haben. Dabei ist zu erkennen, dass eine feinere Diskretisierung nicht zwangsläufig bessere Ergebnisse liefert, da die Stabilitätsprobleme mit feiner werdenden Diskretisierungsgrad zu überwiegen scheinen. Dieser Zusammenhang zwischen feiner werdender Diskretisierung und steigender Instabilität ist typisch für inverse Probleme.

Wir haben noch eine Vielzahl weiterer numerischer Experimente meist mit ähnlichen qualitativen Ergebnissen durchgeführt. Dabei haben wir auch andere Parametervektoren $\underline{p}^* \in D_{\max}$ untersucht. Viele Parameterkonstellationen zeigen ähnliche Inkorrektkeitsphänomene, wie wir sie in der obigen Fallstudie gesehen haben. Es gab allerdings auch einige Fälle von Parametervektoren \underline{p}^* , bei denen wir durch die einprozentige Datenstörung keine negativen Auswirkungen auf die Stabilität beobachten konnten.

Insgesamt scheint wegen der auftretenden Inkorrektkeitsphänomene eine Regularisierung für die Parameteridentifikation erforderlich zu sein. Einen erfolgversprechenden Regularisierungszugang zur stabilen numerischen Lösung von (5.3) stellen wir im folgenden Abschnitt vor.

Kapitel 6

Ein Multiparameter- Regularisierungszugang

Die Instabilität der Lösung \underline{p}^δ des Least-Squares Problems (5.3) bezüglich leicht gestörter Daten \underline{z}^δ , die wir anhand einer numerischen Fallstudie im vorherigen Abschnitt illustriert haben, motiviert die Verwendung eines Regularisierungszuganges. Leider haben wir keine Apriori-Informationen, mit deren Hilfe wir spezielle Parametervektoren favorisieren könnten. Bei der Auswertung der numerischen Fallstudien haben wir allerdings festgestellt, dass die jährlichen Semiinvarianten $s_{1,k}(\underline{p}^\delta) = \frac{1}{\tau} s_{\tau,k}(\underline{p}^\delta)$ der instabilen Lösung \underline{p}^δ deutlich von den Semiinvarianten $s_{1,k}(\underline{p}^*)$ abweichen können, obwohl zwischen den diskretisierten Dichtefunktionen $\tilde{f}_n(\cdot, \underline{p}^\delta)$ und $\tilde{f}_n(\cdot, \underline{p}^*)$ kaum Unterschiede festzustellen waren.

6.1 Zusatzinformation durch Semiinvarianten

Wir nutzen die beobachteten Abweichungen zwischen den Semiinvarianten zur Stabilisierung des inversen Problems. Da die Größenordnung der verschiedenen Semiinvarianten, $s_{1,1}(\underline{p}^\delta)$, $s_{1,2}(\underline{p}^\delta)$, \dots , sich stark unterscheiden, verwenden wir eine Multiparameter-Regularisierung, wie sie in [36, §4.2] eingeführt wurde. Parametervektoren $\underline{p} \in D_{\max}$, deren ersten $k = 1, \dots, l$ Semiinvarianten $s_{1,k}(\underline{p})$ stark von den entsprechenden empirischen Semiinvarianten $s_{1,k}^\delta$ des Datenvektors \underline{z}^δ abweichen, können keine gute Approximation der exakten Lösung \underline{p}^* unseres inversen Problems darstellen¹. Daher regulieren wir das diskretisierte inverse Problem (5.2), indem wir für das Least-Squares Problem (5.3) nur Parametervektoren zulassen, deren Semiinvarianten sich in der Umgebung der empirischen Semiinvarianten befinden. Die Einschränkung des zulässigen

¹Die Semiinvarianten $s_{\tau,k}(\underline{p})$ hängen für alle $k = 1, \dots, l$ stetig von $\underline{p} \in D$ ab. Genauer existiert für jedes $\varepsilon > 0$ ein $\delta > 0$, so dass für Parametervektoren \underline{p} , die sich in einer δ -Umgebung des Parametervektors \underline{p}_0 befinden, wegen (2.19) die Beziehung $|s_{\tau,k}(\underline{p}) - s_{\tau,k}(\underline{p}_0)| \leq \varepsilon$ erfüllt ist. Eine gute Approximation \underline{p}^δ des Parametervektors \underline{p}^* impliziert auch eine gute Approximation von $s_{\tau,k}(\underline{p}^*)$ durch $s_{\tau,k}(\underline{p}^\delta)$.

Bereichs D_{\max} führt zu dem folgenden im Vergleich zu (5.3) stärker restringierten Minimierungsproblem.

Definition 6.1.1 *Wir betrachten als Optimallösung $\underline{p}_{\text{opt}}^\delta$ die Lösung des Extremalproblems*

$$\left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2 \longrightarrow \min \quad \text{für } \underline{p} \in M^\delta \quad (6.1)$$

mit dem eingeschränkten zulässigen Bereich

$$M^\delta := \{ \underline{p} \in D_{\max} : |s_{1,k}(\underline{p}) - s_{1,k}^\delta| \leq \delta_k, \quad k = 1, \dots, l \},$$

wobei \tilde{A} den im vorherigen Kapitel eingeführten diskretisierten Operator (5.2) mit

$$\left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2 = \sum_{i=0}^{n-1} \left(\tilde{f}_n(\tilde{x}_i, \underline{p}) - z_i^\delta \right)^2 \quad (6.2)$$

bezeichnet.

Das Optimierungsproblem (6.1) ist bezüglich der Zielfunktion und bezüglich der Restriktionsmenge M^δ nicht notwendigerweise konvex, so dass wir im folgenden Abschnitt dieses numerisch schwer zu behandelnde restringierte Optimierungsproblem in ein einfacheres mit Box-Constraints $\underline{p} \in D_{\max}$ überführen werden. Ferner werden wir einen Algorithmus zur numerischen Bestimmung der Lösung $\underline{p}_{\text{opt}}^\delta$ vorstellen und anhand einer numerischen Fallstudie die Stabilisierung des inversen Problems nachweisen.

An dieser Stelle wollen wir nochmals kurz verdeutlichen, warum wir die Semiinvarianten und nicht die Momente zur Regularisierung verwenden. Wir nutzen empirische Returns mit Zeitlag τ , um die Parameter des Jump-Diffusions-Prozesses zu schätzen. Aus den empirischen Returns erhalten wir die empirischen Momente $m_{\tau,k}^\delta$ sowie die empirischen Semiinvarianten $s_{\tau,k}^\delta$. Verwenden wir beispielsweise Tageskurse, so haben die ersten Momente und Semiinvarianten sehr kleine Werte. Nach Lemma 2.2.10 gilt aber $\tau s_{1,k}^\delta = s_{\tau,k}^\delta$, so dass wir auch mit den günstiger skalierten jährlichen Semiinvarianten in (6.4) rechnen können. Die Momente erfüllen diese Skalierungseigenschaft nicht, so dass wir dann aus täglichen Returns keine jährlichen Werte erhalten würden.

In Kapitel 4 haben wir den Abstand zwischen der Datenfunktion f^δ und der uns unbekannten exakten Dichtefunktion f^* mit dem Datenfehlerniveau δ abgeschätzt. In praktischen Anwendungen ist die Apriori-Bestimmung des Datenfehlerniveaus δ mit $\|f^* - f^\delta\|_{L^2(\mathbb{R})} \leq \delta$, wie in vielen praktischen inversen Aufgabenstellungen, ein diffiziles Problem. Für den hier gewählten Regularisierungszugang benötigen wir keine direkte Aussage zum Datenfehlerniveau δ , sondern zu den Fehlerschranken $\delta_1, \dots, \delta_l$. Diese können wir zumindest theoretisch mit Hilfe von δ nach oben abschätzen. Gleichung (2.19) stellt den allgemeinen Zusammenhang zwischen den Parametervektoren $\underline{p} \in D$ und den zugehörigen Semiinvarianten $s_{\tau,k}(\underline{p})$ dar. Wegen (2.17) und (2.18) besteht zwischen der exakten Dichtefunktion f^* mit $\tilde{A}(\underline{p}^*) = f^*$ und den Semiinvarianten $s_{\tau,k}^* := s_{\tau,k}(\underline{p}^*)$ ($\underline{p}^* \in U(f^*)$) ebenfalls ein unmittelbarer Zusammenhang. Wir haben jedoch in Abschnitt 4.2 gezeigt, dass für Parametervektoren $\underline{p} \in D \setminus \hat{D}$ der Operator A

nicht injektiv ist, so dass verschiedene Parametervektoren auf die gleiche Dichtefunktion abgebildet werden können. Dies führt zu der Frage, ob eine Zuordnung verschiedener Semiinvarianten zu einer Dichtefunktion möglich ist. Allerdings können für jeden Nichtinjektivitätsfall $A(\underline{p}_1) = A(\underline{p}_2)$ mit $\underline{p}_1 \neq \underline{p}_2$ die Semiinvarianten $s_{\tau,k}(\underline{p}_1)$ und $s_{\tau,k}(\underline{p}_2)$ nicht verschieden sein.

Lemma 6.1.2 *Die Semiinvarianten $s_{1,k}^*$ ($k \in \mathbb{N}$) sind eindeutig durch die Dichtefunktion $f^* \in A(D_{\max})$ bestimmt.*

Beweis. Es sei $\underline{p}^* \in D_{\max}$ eine Lösung von $A(\underline{p}^*) = f^*$. Falls $f^* \in A(\hat{D} \cap D_{\max})$, so ist $\underline{p}^* \in \hat{D} \cap D_{\max}$ wegen Folgerung 4.2.4 und Folgerung 4.2.6 eindeutig. Somit sind auch die Semiinvarianten $s_{1,k}^* = s_{1,k}(\underline{p}^*)$ eindeutig bestimmt. Ist andernfalls $f^* \in A((D \setminus \hat{D}) \cap D_{\max})$, so ist \underline{p}^* wegen Folgerung 4.2.6 nicht eindeutig bestimmt. Mit Lemma 4.2.5 folgt aber, dass alle Diffusionsparameter μ^* und σ^* der Parametervektoren $\underline{p}^* \in U(f^*)$ übereinstimmen. Ferner erfüllen die Sprungparameter immer die Beziehung $\lambda^*((\mu_Y^*)^2 + (\sigma_Y^*)^2) = 0$, so dass mit (2.19) für zwei \underline{p}_1^* und $\underline{p}_2^* \in U(f^*) \cap D_{\max}$ die Semiinvarianten $s_{\tau,k}(\underline{p}_1^*) = s_{\tau,k}(\underline{p}_2^*) = s_{\tau,k}^*$ übereinstimmen. Folglich sind die Semiinvarianten $s_{\tau,k}^*$ eindeutig durch f^* bestimmt \square

Aus dem unmittelbaren Zusammenhang zwischen der exakten Dichtefunktion f^* und den exakten Semiinvarianten $s_{1,k}^*$ folgt aus $\|f^* - f^\delta\|_{L^2(\mathbb{R})} \leq \delta$ für jedes $k = 1, \dots, l$ die Existenz einer Fehlerschranke $\tilde{\delta}_k = \tilde{\delta}_k(\delta)$ mit $|s_{1,k}^* - s_{1,k}^\delta| = \tilde{\delta}_k$. Folglich dürfen wir wegen

$$|s_{1,k}(\underline{p}) - s_{1,k}^\delta| \leq |s_{1,k}(\underline{p}) - s_{1,k}^*| + |s_{1,k}^* - s_{1,k}^\delta| = \tilde{\delta}_k + |s_{1,k}(\underline{p}) - s_{1,k}^*|$$

die Fehlerschranken δ_k nicht zu klein wählen. Würden wir beispielsweise für mindestens ein k die Fehlerschranke δ_k mit $\delta_k < \tilde{\delta}_k$ wählen, so würde $\underline{p}^* \notin M^\delta$ gelten und wir würden somit a priori die exakte Lösung \underline{p}^* als mögliche Lösung ausschließen. Wählen wir andererseits die Fehlerschranken δ_k zu groß, so umfasst die zu den empirischen Daten kompatible Menge M^δ zu viele Parametervektoren \underline{p} , um eine Stabilisierung des instabilen Least-Squares Problems (5.3) zu erreichen.

6.2 Regularisierungsalgorithmus

Zur Bestimmung der Optimallösung $\underline{p}_{opt}^\delta$ aus Definition 6.1.1, die durch Minimierung des Least-Squares Terms $\|\tilde{A}(\underline{p}) - \underline{z}^\delta\|^2$ für Parametervektoren \underline{p} aus der zu den Daten \underline{z}^δ kompatiblen Menge M^δ geschieht, müssen wir ein nicht notwendigerweise konvexes Optimierungsproblem lösen. Dies tun wir mit Hilfe einer Lagrange-Technik, bei welcher wir den Least-Squares Term um den Strafterm

$$\sum_{k=1}^l \alpha_k \left(|s_{1,k}(\underline{p}) - s_{1,k}^\delta|^2 - \delta_k^2 \right)$$

mit dem Lagrange-Multiplikatorenvektor $\underline{\alpha} \in \mathbb{R}_+^l$ ergänzen und das Lagrange-Funktional von Problem (6.1)

$$L(\underline{p}, \underline{\alpha}) = \left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2 + \sum_{k=1}^l \alpha_k \left(|s_{1,k}(\underline{p}) - s_{1,k}^\delta|^2 - \delta_k^2 \right) \quad (6.3)$$

minimieren. Wir werden diesen Vektor $\underline{\alpha} \in \mathbb{R}_+^l$ im Weiteren Regularisierungsparametervektor nennen, da diese Technik einer Multiparameter-Regularisierung, bei der der Least-Squares Term (6.2) um das Regularisierungsfunktional

$$\Omega(\underline{p}, \underline{\alpha}, \underline{z}^\delta) := \sum_{k=1}^l \alpha_k |s_{1,k}(\underline{p}) - s_{1,k}^\delta|^2 \quad (6.4)$$

ergänzt wird, entspricht.

Definition 6.2.1 Wir bezeichnen mit $\underline{p}_{\underline{\alpha}}^\delta$ die mit dem Regularisierungsparametervektor $\underline{\alpha} \in \mathbb{R}_+^l$ regularisierte Lösung des Extremalproblems

$$F_{\text{mult}}(\underline{p}, \underline{\alpha}) = \left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2 + \Omega(\underline{p}, \underline{\alpha}, \underline{z}^\delta) \longrightarrow \min \quad \text{für } \underline{p} \in D_{\text{max}}. \quad (6.5)$$

Wegen (6.4) und (6.2) gilt für das Regularisierungsfunktional

$$\begin{aligned} F_{\text{mult}}(\underline{p}, \underline{\alpha}) &= \left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2 + \Omega(\underline{p}, \underline{\alpha}, \underline{z}^\delta) \\ &= \sum_{i=0}^{n-1} \left(\tilde{f}_n(\tilde{x}_i, \underline{p}) - z_i^\delta \right)^2 + \sum_{k=1}^l \alpha_k |s_{1,k}(\underline{p}) - s_{1,k}^\delta|^2. \end{aligned} \quad (6.6)$$

Durch die Berechnung regularisierter Lösungen $\underline{p}_{\underline{\alpha}}^\delta$ für verschiedene Regularisierungsparametervektoren $\underline{\alpha}$ können wir näherungsweise die Lösung $\underline{p}_{\text{opt}}^\delta$ aus Definition 6.1.1 als Approximation des exakten Parametervektors \underline{p}^* berechnen.

Bemerkung 6.2.2 Das Hinzufügen des Regularisierungsfunktionals Ω , welches vom Regularisierungsparametervektor $\underline{\alpha} \in \mathbb{R}_+^l$ abhängt, erinnert an eine verallgemeinerte Tichonov-Regularisierung (siehe beispielsweise [36]), bei der ein skalarer Regularisierungsparameter den Strafterm Ω wichtet². Allerdings ist das von uns gewählte Regularisierungsfunktional im Gegensatz zu einer herkömmlichen Tichonov-Regularisierung von den Daten \underline{z}^δ abhängig. Bei der Tichonov-Regularisierung verursacht ein zu groß gewählter Regularisierungsparameter auch bei exakten Daten einen Approximationsfehler. Falls wir hinreichend viele exakte Daten ($\delta = 0$) verwenden³, gilt $s_{1,k}(\underline{p}^*) = s_{1,k}^\delta$, so

²Bei der Identifizierung einer Funktion $x \in X$ favorisiert man häufig normkleinste oder glatte Lösungen, so dass bei der Tichonov-Regularisierung häufig $\Omega(x) = \|x\|_X$ oder $\Omega(x) = \|x'\|_X$ gewählt wird.

³Falls wir hinreichend viele exakte Daten verwenden, stimmen die empirische Dichtefunktion f_n^δ und f_{r_τ} überein, sofern der Durchmesser der Diskretisierung gegen null konvergiert. Ferner sind für $\delta = 0$ die in Kapitel 5 eingeführte zufällig gestörte Funktion f_n^δ und die exakte diskretisierte Dichtefunktion f_n^* identisch. Für $n \rightarrow \infty$ und $\Delta \hat{x} = \Delta \hat{x} \rightarrow 0$ stimmen daher die Semiinvarianten und empirischen Semiinvarianten, die über die empirischen Momente $m_{1,k}^\delta$ (vgl. Formel (5.4)) berechnet werden, überein.

dass der Regularisierungsterm $\Omega(\underline{p}^*, \underline{\alpha}, \underline{z}^\delta)$ für jeden Regularisierungsparametervektor $\underline{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_l)^T \in \mathbb{R}_+^l$ verschwindet. Folglich verursacht der Strafterm (6.4) bei der Verwendung von (hinreichend vielen) exakten Daten für jeden beliebigen Regularisierungsparametervektor $\underline{\alpha} \in \mathbb{R}_+^l$ keinen Approximationsfehler, d.h. es gilt $\underline{p}^* = \underline{p}_{\underline{\alpha}}^\delta$. Da in praktischen Fällen aber die Schätzung der empirischen Momente bzw. Semiinvarianten aus relativ wenig Datensätzen nicht sehr zuverlässig ist, kann für zu große Regularisierungsparameter α_k der Regularisierungsterm $\Omega(\underline{p}, \underline{\alpha}, \underline{z}^\delta)$ übergewichtet werden und folglich ist es möglich, dass die Lösung $\underline{p}_{\underline{\alpha}}^\delta$ keine gute Approximation des exakten Parametervektors \underline{p}^* darstellt. Wählen wir andererseits zu kleine Regularisierungsparameter α_k , nimmt der Regularisierungsterm für alle Parametervektoren $\underline{p} \in D_{\max}$ kleine Werte an, so dass er kaum Wirkung zeigt und folglich die Instabilitätsprobleme dominieren. Folglich muss für praktisch verwertbare Lösungen ein Kompromiss zwischen zu kleinen und zu großen Regularisierungsparametern α_k ($k = 1, \dots, l$) gefunden werden.

Wir werden im Folgenden den Zusammenhang zwischen $\underline{p}_{\underline{\alpha}}^\delta$ und $\underline{p}_{\text{opt}}^\delta$ näher untersuchen. Die Lösung $\underline{p}_{\text{opt}}^\delta$ aus Definition 6.1.1 wird durch die Apriori-Fehlerschranken $\delta_1, \dots, \delta_l$ bestimmt, während die Lösung $\underline{p}_{\underline{\alpha}}^\delta$ aus Definition 6.2.1 vom Regularisierungsparametervektor $\underline{\alpha}$ abhängig ist. Dieser Zusammenhang zwischen beiden Problemen wird auch bei dem so genannten Diskrepanzprinzip genutzt (siehe z.B. [38]), für das aus dem Datenfehler δ der optimale Regularisierungsparameter für die Tichonov-Regularisierung berechnet wird.

Das Lagrange-Funktional (6.3) lässt sich zu

$$L(\underline{p}, \underline{\alpha}) = F_{\text{mult}}(\underline{p}, \underline{\alpha}) - \sum_{k=1}^l \alpha_k \delta_k^2 \quad (6.7)$$

umformen. Theorem 4.22 aus [36] stellt den Zusammenhang zwischen dem Lagrangeschen Optimierungsproblem (6.1) und dem regularisierten Problem (6.5) dar. Wir formulieren und beweisen hier diesen Satz für unsere Zwecke.

Satz 6.2.3 *Das Paar $(\hat{\underline{p}}, \hat{\underline{\alpha}}) \in D_{\max} \times \mathbb{R}_+^l$ ist ein Sattelpunkt des Lagrangeschen-Sattelpunktproblems (6.1), genau dann wenn $\hat{\underline{p}} = \underline{p}_{\hat{\underline{\alpha}}}^\delta$ eine mit dem Regularisierungsparametervektor $\hat{\underline{\alpha}}$ regularisierte Lösung von (6.5) ist, die die Bedingungen*

$$\hat{\alpha}_k \left(|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 - \delta_k^2 \right) = 0 \quad \text{für alle } k = 1, \dots, l \quad (6.8)$$

und

$$|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 \leq \delta_k^2 \quad \text{falls } \hat{\alpha}_k = 0 \quad \text{für alle } k = 1, \dots, l \quad (6.9)$$

erfüllt. Jeder Parametervektor $\hat{\underline{p}} = \underline{p}_{\hat{\underline{\alpha}}}^\delta$, der Lösung von (6.5) ist und die Bedingungen (6.8) sowie (6.9) erfüllt, ist eine Lösung $\underline{p}_{\text{opt}}^\delta$ des Optimierungsproblems (6.1).

Beweis. Sei zunächst $(\hat{\underline{p}}, \hat{\underline{\alpha}})$ Sattelpunkt des Lagrangeschen-Sattelpunktproblems (6.1). Dann gilt für alle $\underline{p} \in D_{\max}$ und $\underline{\alpha} \in \mathbb{R}_+^l$ die Beziehung

$$L(\hat{\underline{p}}, \underline{\alpha}) \leq L(\hat{\underline{p}}, \hat{\underline{\alpha}}) \leq L(\underline{p}, \hat{\underline{\alpha}}). \quad (6.10)$$

Aus der zweiten Ungleichung von (6.10) erhalten wir für alle $\underline{p} \in D_{\max}$ mit (6.7) unmittelbar

$$F_{\text{mult}}(\hat{\underline{p}}, \hat{\underline{\alpha}}) \leq F_{\text{mult}}(\underline{p}, \hat{\underline{\alpha}}) \quad \forall \underline{p} \in D_{\max},$$

so dass $\hat{\underline{p}}$ Minimum von (6.5) ist. Ferner folgt mit (6.6) und (6.7) aus der Ungleichung $L(\hat{\underline{p}}, \underline{\alpha}) \leq L(\hat{\underline{p}}, \hat{\underline{\alpha}})$ von (6.10) die Beziehung

$$\Omega(\hat{\underline{p}}, \underline{\alpha}, \underline{z}^\delta) - \Omega(\hat{\underline{p}}, \hat{\underline{\alpha}}, \underline{z}^\delta) = F_{\text{mult}}(\hat{\underline{p}}, \underline{\alpha}) - F_{\text{mult}}(\hat{\underline{p}}, \hat{\underline{\alpha}}) \leq \sum_{k=1}^l \delta_k^2 (\alpha_k - \hat{\alpha}_k).$$

Dies impliziert wegen

$$\Omega(\hat{\underline{p}}, \underline{\alpha}, \underline{z}^\delta) - \Omega(\hat{\underline{p}}, \hat{\underline{\alpha}}, \underline{z}^\delta) = \sum_{k=1}^l (\alpha_k - \hat{\alpha}_k) |s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2$$

für alle $\underline{\alpha} \in \mathbb{R}_+^l$

$$\sum_{k=1}^l (\hat{\alpha}_k - \alpha_k) \left(|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 - \delta_k^2 \right) \geq 0.$$

Da diese Ungleichung für alle Vektoren $\underline{\alpha} \in \mathbb{R}_+^l$ erfüllt ist, muss für jedes $k = 1, \dots, l$ und jedes $\alpha_k \geq 0$ die Beziehung

$$(\hat{\alpha}_k - \alpha_k) \left(|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 - \delta_k^2 \right) \geq 0$$

gelten, woraus unmittelbar

$$|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 - \delta_k^2 = 0$$

folgt. Falls $\hat{\alpha}_k = 0$ gilt, muss wegen $\alpha_k \geq 0$ die Beziehung $|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 - \delta_k^2 \leq 0$ erfüllt sein, so dass (6.8) und (6.9) gezeigt sind.

Sei nun $\hat{\underline{p}} = \underline{p}_{\hat{\underline{\alpha}}}$ Lösung von (6.5), die (6.8) und (6.9) erfüllt. Wegen (6.8) und (6.9) ist $\hat{\underline{p}} = \underline{p}_{\hat{\underline{\alpha}}}$ aus der Menge M^δ . Außerdem gilt

$$F_{\text{mult}}(\hat{\underline{p}}, \hat{\underline{\alpha}}) \leq F_{\text{mult}}(\underline{p}, \hat{\underline{\alpha}})$$

für alle Parametervektoren $\underline{p} \in D_{\max}$. Somit folgt mit (6.6) für alle Parametervektoren $\underline{p} \in M^\delta$ wegen (6.9) die Beziehung

$$\begin{aligned} \left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2 - \left\| \tilde{A}(\hat{\underline{p}}) - \underline{z}^\delta \right\|^2 &\geq \Omega(\hat{\underline{p}}, \hat{\underline{\alpha}}, \underline{z}^\delta) - \Omega(\underline{p}, \hat{\underline{\alpha}}, \underline{z}^\delta) \\ &= \sum_{k=1}^l \hat{\alpha}_k \left(|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 - |s_{1,k}(\underline{p}) - s_{1,k}^\delta|^2 \right) \\ &\geq \sum_{k=1}^l \hat{\alpha}_k \left(|s_{1,k}(\hat{\underline{p}}) - s_{1,k}^\delta|^2 - \delta_k^2 \right) = 0, \end{aligned}$$

d.h. $\hat{p} = \underline{p}_{\hat{\alpha}}$ ist eine Lösung des Lagrangeschen Problems (6.1), da $\left\| \tilde{A}(\hat{p}) - \underline{z}^\delta \right\|^2 \leq \left\| \tilde{A}(\underline{p}) - \underline{z}^\delta \right\|^2$ für alle $\underline{p} \in M^\delta$ gilt. Damit ist obige Aussage bewiesen \square

Falls ein Sattelpunkt $(\hat{p}, \hat{\alpha}) \in D_{\max} \times \mathbb{R}_+^l$ im Sinne von Satz 6.2.3 existiert, so ist $\hat{p} = \underline{p}_{opt}^\delta$ eine mit dem Regularisierungsparametervektor $\hat{\alpha}$ regularisierte Lösung von (6.1). Daher können wir anstelle des nichtkonvexen Optimierungsproblems (6.1) auch das einfachere Problem⁴ (6.5) lösen. Der benötigte optimale Regularisierungsparametervektor $\hat{\alpha}$ lässt sich aus den Beziehungen (6.8) und (6.9) bestimmen. Deshalb betrachten wir nun einen iterativen Lösungsalgorithmus, um approximativ einen Sattelpunkt zu berechnen.

6.2.1 Eine Fixpunktiteration

Wir können die Bedingung (6.8) als Fixpunktgleichung

$$\hat{\alpha}_k = \hat{\alpha}_k \frac{|s_{1,k}(\hat{p}) - s_{1,k}^\delta|^2}{\delta_k^2} \quad \text{für alle } k = 1, \dots, l$$

umformulieren und nutzen diese für eine Fixpunktiteration. Dafür wählen wir ein kleines, positives ε mit $0 < \varepsilon \ll 1$ und einen Startvektor $\underline{\alpha}^{(0)} \in \mathbb{R}_+^l$ mit positiven Komponenten $\alpha_k^{(0)} > 0$ ($k = 1, \dots, l$) und starten die Iteration

$$\begin{aligned} \underline{p}^{(j)} &:= \underline{p}_{\underline{\alpha}^{(j)}}^\delta \quad \text{für } j = 0, 1, \dots; \\ \alpha_k^{(j+1)} &:= \alpha_k^{(j)} \max \left\{ \frac{|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^2}{\delta_k^2}, \varepsilon \right\} \quad \text{für } j = 0, 1, \dots; k = 1, \dots, l. \end{aligned} \tag{6.11}$$

In Analogie zu Theorem 4.24 aus [36] erhalten wir den folgenden Satz, dessen Beweis wir hier präzisieren.

Satz 6.2.4 *Falls die Iteration (6.11) konvergiert, d.h. $\underline{\alpha}^{(j)} \rightarrow \hat{\alpha} \in \mathbb{R}_+^l$ und $\underline{p}^{(j)} \rightarrow \hat{p} \in D_{\max}$ für $j \rightarrow \infty$ gelten, ist das Paar $(\hat{p}, \hat{\alpha})$ Sattelpunkt des Lagrangeschen-Sattelpunkt Problems (6.1) und es gilt $\hat{p} = \underline{p}_{\hat{\alpha}}^\delta$, d.h. \hat{p} ist eine optimale Lösung von (6.5).*

Beweis. Da die Folge der Regularisierungsparametervektoren $\{\underline{\alpha}^{(j)}\}_j$ der Iteration (6.11) gegen $\hat{\alpha}$ und die zugehörige Folge der Lösungen $\{\underline{p}^{(j)}\}_j = \{\underline{p}_{\underline{\alpha}^{(j)}}^\delta\}_j$ gegen \hat{p} konvergieren, ist für alle $\underline{p} \in D_{\max}$ die Beziehung

$$F_{\text{mult}}(\hat{p}, \hat{\alpha}) \leq F_{\text{mult}}(\underline{p}, \hat{\alpha}) \tag{6.12}$$

⁴Das Optimierungsproblem (6.5) ist ein nichtlineares Least-Squares Problem mit der einfachen Nebenbedingung $\underline{p} \in D_{\max}$ (Box-Constraints), für das gute Lösungsalgorithmen (z.B. Matlab-Routine `lsqcurvefit`) existieren.

sowie für alle $k = 1, \dots, l$ die Beziehung

$$\hat{\alpha}_k = \hat{\alpha}_k \max \left\{ \frac{|s_{1,k}(\hat{p}) - s_{1,k}^\delta|^2}{\delta_k^2}, \varepsilon \right\} \quad (6.13)$$

erfüllt. Wegen (6.12) ist \hat{p} Lösung von (6.5). Gleichung (6.13) ist für $\hat{\alpha}_k = 0$ oder $|s_{1,k}(\hat{p}) - s_{1,k}^\delta|^2 = \delta_k^2$ erfüllt, womit (6.8) gilt. Es bleibt noch zu zeigen, dass für $\hat{\alpha}_k = 0$ die Beziehung $|s_{1,k}(\hat{p}) - s_{1,k}^\delta|^2 \leq \delta_k^2$ erfüllt ist. Sei dazu $|s_{1,k}(\hat{p}) - s_{1,k}^\delta|^2 > \delta_k^2$. Dann existiert ein j_0 , so dass $|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^2 > \delta_k^2$ für alle $j \geq j_0$ gilt, da die k -te Semiinvariante für alle Parametervektoren $\underline{p} \in D$ stetig ist. Folglich gilt

$$\max \left\{ \frac{|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^2}{\delta_k^2}, \varepsilon \right\} > 1,$$

womit sich aus der Transformation (6.11) für alle $j \geq j_0$ unmittelbar $\alpha_k^{(j+1)} > \alpha_k^{(j)}$ ergibt. Da alle Komponenten des Startvektors $\underline{\alpha}^0$ positiv sind, d.h. $\alpha_k^{(0)} > 0$, sowie $0 < \varepsilon \ll 1$ gilt, folgt aus der Transformation (6.11) unmittelbar $\alpha_k^{(j)} > 0$ für jedes j . Somit gilt $\alpha_k^{(j+1)} > \alpha_k^{(j)} > 0$, was aber einen Widerspruch zur Konvergenz der Folge $\{\alpha_k^{(j)}\}_j$ gegen $\hat{\alpha}_k = 0$ darstellt. Daher muss $|s_{1,k}(\hat{p}) - s_{1,k}^\delta|^2 \leq \delta_k^2$ für $\hat{\alpha}_k = 0$ gelten, so dass auch (6.9) erfüllt ist. Die optimale Lösung \hat{p} von (6.5) erfüllt somit (6.8) und (6.9), weshalb mit Satz 6.2.3 dann $(\hat{p}, \hat{\alpha})$ Sattelpunkt des Lagrangeschen-Sattelpunkt Problems (6.1) ist \square

Die Angabe hinreichender Bedingungen für die in Satz 6.2.4 vorausgesetzte Konvergenz der Folge $\{\underline{\alpha}^{(j)}\}_j$ gegen $\hat{\alpha}$ und der Folge $\{\underline{p}^{(j)}\}_j$ gegen \hat{p} für $j \rightarrow \infty$ ist wohl recht schwierig. Es sind uns keine Aussagen zu Bedingungen bekannt, die eine Kontraktion im Sinne des Banachschen Fixpunktsatzes sichern. Allerdings konvergiert der folgende Algorithmus in vielen praktischen Situationen gegen akzeptable Lösungen, die \underline{p}^* gut approximieren, sofern die Schranken $\delta_1, \dots, \delta_l$ der Semiinvarianten vernünftig gewählt sind.

Algorithmus 6.2.5

- 0. Schritt** Wähle $j_{\max} \in \mathbb{N}$, kleine Werte $\varepsilon, \varepsilon_1 > 0$ und einen Startvektor $\underline{\alpha}^{(0)} \in \mathbb{R}_+^k$. Setze $j := 0$.
- 1. Schritt** Bestimme $\underline{p}^{(j)} = \underline{p}_{\underline{\alpha}^{(j)}}^\delta$ als Minimum von $F_{\text{mult}}(\underline{p}, \underline{\alpha}^{(j)})$, d.h. bestimme für den Regularisierungsvektor $\underline{\alpha}^{(j)}$ die Lösung des Minimierungsproblems (6.5).
- 2. Schritt** Berechne $\alpha_k^{(j+1)} = \alpha_k^{(j)} \max \left\{ \frac{|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^2}{\delta_k^2}, \varepsilon \right\} \quad (k = 1, 2, \dots, l)$.
- 3. Schritt** Falls $\|\underline{\alpha}^{(j+1)} - \underline{\alpha}^{(j)}\| \leq \varepsilon_1$ oder $j+1 \geq j_{\max}$ setze $\hat{p} := \underline{p}^{(j)}$ und stoppe; setze andernfalls $j := j+1$ und fahre mit dem 1. Schritt fort.

Bemerkung 6.2.6 Wir haben für jeden Iterationsschritt das Optimierungsproblem (6.5) numerisch zu lösen. Als Startvektor $\underline{p}_0^{(j)}$ der j -ten Iteration wählen wir mit $\underline{p}^{(j-1)}$ die Lösung der vorherigen Iteration.

6.2.2 Numerisches Beispiel

Wir haben die Zuverlässigkeit der Multiparameter-Regularisierung anhand einer umfangreichen numerischen Fallstudie überprüft. In diesem Abschnitt wollen wir ein repräsentatives numerisches Beispiel zur Multiparameter-Regularisierung vorstellen. Dazu haben wir in Analogie zu dem Beispiel aus Abschnitt 5.3 für den Parametervektor $\underline{p}^* = (0.1, 0.2, 10, 0.1, 0.2)$ den exakten Datenvektor \underline{z}^* mit dem Datenfehlniveau δ zufällig gestört und mit dem Algorithmus 6.2.5 die optimale regularisierte Lösung $\hat{\underline{p}} = \underline{p}_{\hat{\alpha}}^\delta$ aus den Daten \underline{z}^δ bestimmt. Im Gegensatz zu der Fallstudie aus dem vorherigen Kapitel, die zur Illustration der Inkorrekttheitsphänomene diente und bei der wir die Daten um ein Prozent gestört hatten, verwenden wir hier eine 10-prozentige Datenstörung, da dies einer realistischeren Situation entspricht. Die restlichen Einstellungen, wie beispielsweise die Diskretisierungsparameter⁵ \tilde{x}_0 , \tilde{x}_n und n haben wir, wie in Tabelle 6.1 zu sehen ist, analog zur oben diskutierten Fallstudie des vorherigen Kapitels gewählt. Ferner haben wir aus den gleichen Gründen wie bei der Fallstudie zur Illustration der Inkorrekttheitsphänomene $D_{\max} = D$ gesetzt.

μ	σ	λ	μ_Y	σ_Y	τ	δ	n	\tilde{x}_0	\tilde{x}_n	l	ε	ε_1
0.1	0.2	10	0.1	0.2	0.004	0.1	40	-0.1	0.1	5	1e-05	1e-05

Tabelle 6.1: Werte des Parametervektors \underline{p}^* und weitere Einstellungen für die Multiparameter-Regularisierung

Zur Regularisierung verwenden wir in diesem Beispiel die ersten fünf Semiinvarianten, d.h. wir wählen $l = 5$, so dass wir für die Multiparameter-Regularisierung vorher die fünf Schranken $\delta_1, \dots, \delta_5$ festlegen können.

k	Semiinvarianten				Multipar.-Reg.	
	$s_{1,k}^\delta$	$s_{1,k}(\underline{p}^*)$	$ s_{1,k}(\underline{p}^*) - s_{1,k}^\delta $	Abw.	δ_k	$\alpha_k^{(0)}$
1	-0.17002795	-0.19496852	0.02494	12.79%	0.01	10.0
2	+0.54421567	+0.54000000	0.00422	0.78%	0.01	10.0
3	+0.13101847	+0.13000000	0.00102	0.78%	0.002	10.0
4	+0.07375075	+0.07300000	0.00075	1.00%	0.002	10.0
5	+0.02837745	+0.02810000	0.00028	0.99%	0.002	10.0

Tabelle 6.2: Vergleich der Semiinvarianten und entsprechende Einstellungen für die Multiparameter-Regularisierung

⁵Wir benötigen zur Berechnung der empirischen Momente $m_{\tau,k}^\delta$, wie in Abschnitt 5.3 eingeführt, die Werte \hat{x}_0 und \hat{x}_m . Diese Werte bestimmen wir, wie in Bemerkung 5.3.1 beschrieben, so dass wir diese Werte nicht explizit vorgeben müssen.

In Tabelle 6.2 sind die exakten Semiinvarianten $s_{1,k}(\underline{p}^*)$, die Semiinvarianten der Daten $s_{1,k}^\delta$, deren Abweichungen und die verwendeten Niveauschranken δ_k für die ersten fünf Semiinvarianten abgebildet.

Für die von uns simulierten Daten \underline{z}^δ sowie die oben gewählten Einstellungen konvergiert der Algorithmus zur Multiparameter-Regularisierung sehr gut und liefert, wie in Tabelle 6.3 zu sehen ist, eine zufriedenstellende Approximation $\hat{\underline{p}} = \underline{p}_{\hat{\underline{a}}}^\delta$ des Parametervektors \underline{p}^* .

Parameter	\underline{p}^*	$\hat{\underline{p}} = \underline{p}_{\hat{\underline{a}}}^\delta$	Abweichung
μ	0.1	0.115857	15.86%
σ	0.2	0.197669	1.17%
λ	10.0	9.491907	5.08%
μ_Y	0.1	0.103871	3.87%
σ_Y	0.2	0.203410	1.71%

Tabelle 6.3: Multiparameter-regularisierte Lösung

Obwohl wir über keine hinreichenden Bedingungen verfügen, um die Konvergenz zu sichern, konvergierte der Algorithmus gegen einen Sattelpunkt und die durch (6.11) entstandene Folge $\{\underline{p}^{(j)}\}_j$ wies die für Fixpunktiterationen typische lineare Konvergenz auf.

In Tabelle 6.3 haben wir außerdem die Abweichungen (in Prozent) zwischen den fünf Parametern der regularisierten Lösung und den exakten Parametern dargestellt. Erwartungsgemäß sind die Abweichungen des Diffusionsparameters μ mit ca. 16% am höchsten. Die restlichen Parameter wurden dagegen sehr gut geschätzt. So liegen beispielsweise die Abweichungen von σ und σ_Y sogar unter zwei Prozent. Der optimale Regularisierungsparametervektor $\hat{\underline{a}}$ ist in der letzten Spalte von Tabelle 6.4 zu sehen. Bei zahlreichen weiteren Rechnungen hat sich gezeigt, dass die regularisierten Lösungen nicht mehr so stark von dem gewählten Startvektor \underline{p}_0 abhängen wie im unregularisierten Fall. Tabelle 6.4 illustriert, dass in diesem Beispiel nur für die zweite und dritte

k	$s_{1,k}^\delta$	$s_{1,k}(\underline{p}_{\hat{\underline{a}}}^\delta)$	$ s_{1,k}(\underline{p}_{\hat{\underline{a}}}^\delta) - s_{1,k}^\delta $	δ_k	$\hat{\underline{a}}$
1	-0.17002795	-0.17002795	0.00680543	0.01	0.00000000
2	+0.54421567	+0.53421568	0.00999999	0.01	0.28657429
3	+0.13101847	+0.13301848	0.00200001	0.002	6.34543394
4	+0.07375075	+0.07527707	0.00152632	0.002	0.00000000
5	+0.02837745	+0.02983393	0.00145648	0.002	0.00000000

Tabelle 6.4: Optimale Regularisierungsparameter sowie Vergleich der Apriori- und Aposteriori-Abweichungen

Semiinvariante die Restriktionen der empirischen Semiinvarianten durch δ_2 und δ_3 in der optimalen Lösung aktiv sind.

Ein Vergleich der Semiinvarianten $s_{1,k}(\underline{p}_{\hat{\underline{a}}}^\delta)$ des geschätzten Parametervektors $\hat{\underline{p}} = \underline{p}_{\hat{\underline{a}}}^\delta$ der Multiparameter-Regularisierung mit den Semiinvarianten $s_{1,k}(\underline{p}^*)$ des exakten Parametervektors \underline{p}^* und den Semiinvarianten $s_{1,k}^\delta$ der gestörten Daten zeigt die Anpassung

der Semiinvarianten der Lösung an die empirischen Semiinvarianten. In Tabelle 6.5 ist dieser Zusammenhang illustriert.

k	$s_{1,k}(\underline{p}^*)$	$s_{1,k}^\delta$	$s_{1,k}(\underline{p}_\alpha^\delta)$	Abw. zu $s_{1,k}(\underline{p}^*)$	Abw. zu $s_{1,k}^\delta$
1	-0.19496852	-0.17002795	-0.17683338	9.30%	4.00%
2	+0.54000000	+0.54421567	+0.53421568	1.07%	1.84%
3	+0.13000000	+0.13101847	+0.13301848	2.32%	1.53%
4	+0.07300000	+0.07375075	+0.07527707	3.12%	2.07%
5	+0.02810000	+0.02837745	+0.02983393	6.17%	5.13%

Tabelle 6.5: Vergleich der Semiinvarianten

Zusammenfassend können wir feststellen, dass der Multiparameter-Regularisierungszugang, der die Least-Squares Anpassung an die empirische Dichtefunktion mit einer Anpassung der empirischen Semiinvarianten kombiniert, für die untersuchten Beispiele sehr gut funktioniert. Die vorgestellte Regularisierung kann die in Kapitel 5 gezeigten Instabilitätsprobleme, die bei dem herkömmlichen Least-Squares Zugang bereits bei kleinen Datenfehlern auftreten können, gut kompensieren und führt zu akzeptablen Approximationen des exakten Parametervektors. Voraussetzung für die Konvergenz von Algorithmus 6.2.5 ist allerdings eine vernünftige Apriori-Wahl der Fehlerschranken $\delta_1, \dots, \delta_l$, die in praktischen Situationen unbekannt sind. Die Wahl eines zu kleinen δ_k birgt die Gefahr, dass wenn $|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|$ das δ_k -Niveau deutlich überschreitet wegen

$$\alpha_k^{(j+1)} = \alpha_k^{(j)} \max \left\{ \frac{|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^2}{\delta_k^2}, \varepsilon \right\}$$

der k -te Regularisierungsparameter im $(j+1)$ -ten Iterationsschritt $\alpha_k^{(j+1)}$ sehr große Werte annehmen kann. Folglich dominiert die Anpassung an die k -te empirische Semiinvariante $s_{1,k}^\delta$ die restlichen Ziele, d.h. die Minimierung des Least-Squares Terms (6.2) sowie die Anpassung der restlichen Semiinvarianten an die empirischen Semiinvarianten. Wählen wir andererseits die Fehlerschranken δ_k zu groß, so wird die k -te Restriktion $|s_{1,k}(\underline{p}) - s_{1,k}^\delta| \leq \delta_k$ nie aktiv und der zugehörige Regularisierungsparameter α_k verschwindet, weshalb dann das Ziel der Stabilisierung des instabilen Least-Squares Problems (5.3) nicht mehr erreicht werden kann.

Bei einigen Beispielen der durchgeführten Fallstudie konnten wir bei zu kleiner Wahl der Fehlerschranken $\delta_1, \dots, \delta_l$ feststellen, dass einzelne Regularisierungsparameter α_k tatsächlich sehr große Werte angenommen haben. Trotz alledem veränderten sich im Laufe der Iteration die Werte der Parametervektoren $\underline{p}^{(j)}$ größtenteils nur ab der vierten bzw. fünften Nachkommastelle. Da bei zu kleinem δ_k die k -te Nebenbedingung nahezu immer aktiv ist, schwankt der Wert $\frac{|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^2}{\delta_k^2}$ um die Eins, so dass sich in der $(j+1)$ -ten Iteration der k -ten Regularisierungsparameter $\alpha_k^{(j+1)}$ als Produkt von $\alpha_k^{(j)}$ und einem Wert, der um eins liegt, ergibt. Da $\alpha_k^{(j)}$ bereits einen großen Wert hatte, implizierte dies recht hohe absolute Abweichungen $\|\underline{\alpha}^{(j+1)} - \underline{\alpha}^{(j)}\|$, obwohl für alle k die relativen Veränderungen $\frac{|\alpha_k^{(j+1)} - \alpha_k^{(j)}|}{\alpha_k^{(j)}}$ verhältnismäßig klein waren. Folglich konvergierte

die Folge der Regularisierungsparametervektoren $\{\underline{\alpha}^{(j)}\}_j$ nicht, so dass der Algorithmus 6.2.5 auch nicht abgebrochen hat, obwohl die Folge $\{\underline{p}^{(j)}\}_j$ konvergierte. Diese Erfahrungen motivieren eine modifizierte Abbruchbedingung für den Multiparameter-Regularisierungsalgorithmus, bei der wir die Iteration (6.11) neben der Abbruchbedingung zusätzlich abbrechen, sobald die Abweichungen $\|\underline{p}^{(j+1)} - \underline{p}^{(j)}\|$ unter eine vorgegebene Schranke ε_2 mit $0 < \varepsilon_2 \ll 1$ fallen. Numerische Rechnungen mit $\varepsilon_2 = 0.000001$ haben gezeigt, dass der Algorithmus mit modifizierter Abbruchbedingung gleichwertige Ergebnisse liefert. Das teilweise auftretende Problem, dass bei großen Werten von $\|\underline{\alpha}^{(j+1)} - \underline{\alpha}^{(j)}\|$ der Algorithmus nicht abbrach, obwohl die Folge $\{\underline{p}^{(j)}\}_j$ konvergierte, kann durch diese modifizierte Abbruchbedingung zufriedenstellend gelöst werden. Wir stellen im folgenden Abschnitt eine weitere Modifikation von Algorithmus 6.2.5 zur Verringerung der benötigten Iterationsschritte vor.

6.3 Numerische Modifikation

Der Algorithmus 6.2.5 basiert auf der Fixpunktgleichung

$$\hat{\alpha}_k = \hat{\alpha}_k \frac{|s_{1,k}(\hat{p}) - s_{1,k}^\delta|^2}{\delta_k^2} \quad \text{für alle } k = 1, \dots, l$$

die wir aus der Beziehung (6.8)

$$\hat{\alpha}_k \left(|s_{1,k}(\hat{p}) - s_{1,k}^{\delta_k}|^2 - \delta_k^2 \right) = 0, \quad \forall k = 1, \dots, l,$$

abgeleitet haben. Diese Beziehung ist für jedes $\psi > 0$ äquivalent zu

$$\hat{\alpha}_k \left(|s_{1,k}(\hat{p}) - s_{1,k}^{\delta_k}|^\psi - \delta_k^\psi \right) = 0, \quad \forall k = 1, \dots, l,$$

so dass wir analog zu [37] die Iteration (6.11) mittels

$$\underline{p}^{(j)} = \underline{p}_{\underline{\alpha}^{(j)}}^\delta \quad \text{für } j = 0, 1, \dots;$$

$$\alpha_k^{(j+1)} = \alpha_k^{(j)} \max \left\{ \frac{|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^\psi}{\delta_k^\psi}, \varepsilon \right\} \quad \text{für } j = 0, 1, \dots; k = 1, \dots, l \quad (6.14)$$

modifizieren.

Bemerkung 6.3.1 Die Aussagen von Satz 6.2.4 bleiben für jedes $\psi > 0$ erhalten, d.h. falls die Folge $\{\underline{\alpha}^{(j)}\}_j$ mit $\alpha_k^{(j+1)} = \alpha_k^{(j)} \max \left\{ \frac{|s_{1,k}(\underline{p}^{(j)}) - s_{1,k}^\delta|^\psi}{\delta_k^\psi}, \varepsilon \right\}$ gegen $\hat{\alpha} \in \mathbb{R}_+^l$ und die Folge $\{\underline{p}^{(j)}\}_j$ mit $\underline{p}^{(j)} = \underline{p}_{\underline{\alpha}^{(j)}}^\delta$ für $j \rightarrow \infty$ gegen $\hat{p} \in D_{\max}$ konvergiert, so ist $\hat{p} = \underline{p}_{\hat{\alpha}}^\delta$ eine optimale Lösung von (6.5) und das Paar $(\hat{p}, \hat{\alpha})$ Sattelpunkt des Lagrangeschen-Sattelpunkt Problems (6.1). Folglich beeinflusst ψ nicht die Lösung $(\hat{p}, \hat{\alpha})$, sondern die Konvergenzgeschwindigkeit.

Wir wollen eine kleine numerische Fallstudie zum Einfluss des Exponentialparameters ψ auf den modifizierten Algorithmus zur Multiparameter-Regularisierung, bei dem wir die Regularisierungsparameter durch die Iteration (6.14) berechnen, vorstellen.

6.3.1 Fallstudie zum Einfluss des Exponentialparameters

Um den Einfluss des Exponentialparameters ψ zu analysieren, betrachten wir wieder das Beispiel des vorherigen Abschnittes und lösen Problem (6.1) für den zufällig gestörten Datenvektor \underline{z}^δ mit dem modifizierten Algorithmus jeweils für $\psi = 0.1, \dots, 4.5$. Dabei verwenden wir die Einstellungen aus Tabelle 6.1 und die a priori gewählten Fehler-schranken $\delta_1, \dots, \delta_5$ aus Tabelle 6.2.

Während dieser Fallstudie haben wir die Erfahrung gemacht, dass für $\psi > 2$ der Algorithmus teilweise nicht konvergierte. Genauer ist der im vorherigen Abschnitt beschriebene Effekt aufgetreten, dass die Folge $\{\underline{p}^{(j)}\}_j$ konvergierte, obwohl die Folge $\{\underline{\alpha}^{(j)}\}_j$ nicht konvergierte. Daher haben wir für $\psi > 2$ die maximale Iterationsanzahl j_{\max} aus Effizienzgründen von 5000 auf 3000 herabgesetzt. Allerdings haben wir bei diesem Beispiel zur Berechnung der regularisierten Lösung $\hat{\underline{p}} = \underline{p}_{\hat{\underline{\alpha}}}^\delta$ mittels der modifizierten Transformation (6.14) nur die ursprüngliche Abbruchbedingung $\|\underline{\alpha}^{(j+1)} - \underline{\alpha}^{(j)}\| \leq \varepsilon_1$ berücksichtigt, um einerseits den Einfluss von ψ zu isolieren und andererseits eine bessere Vergleichbarkeit zu dem numerischen Beispiel aus Abschnitt 6.2.2 zu erhalten.

Der Multiparameter-Regularisierungsalgorithmus konvergierte in allen fünf Simulationen für die meisten ψ , insbesondere für $\psi < 2$. Wir haben hier nur die Ergebnisse der ersten (Tabelle 6.6) Simulation dargestellt, da diese repräsentativ für alle durchgeführten Rechnungen sind. Die Tabelle zeigt für jedes ψ die geschätzten Werte der Parametervektoren sowie die Anzahl der benötigten Iterationen. Dabei ist deutlich zu erkennen, dass im Falle der Konvergenz der Exponent ψ keinen Einfluss auf die Lösung $\hat{\underline{p}} = \underline{p}_{\hat{\underline{\alpha}}}^\delta$ hat. In den wesentlichen Nachkommastellen stimmen alle geschätzten Werte überein.

Tabelle 6.7 zeigt allerdings, dass sich der Exponentialparameter deutlich auf die Anzahl der benötigten Iterationen und folglich auf die benötigte Rechenzeit auswirkt. In Abbildung 6.1 sind die benötigten Iterationen nochmals grafisch dargestellt. Für jede der fünf Simulationen haben wir neue Daten generiert und mit diesen die Multiparameter-Regularisierung für alle Exponentialparameter durchgeführt. Sowohl in Tabelle 6.7 als auch in Abbildung 6.1 ist zu erkennen, dass sich die benötigten Iterationszahlen der jeweiligen Simulationen unterscheiden. Die zufälligen Störungen des Datenvektors \underline{z}^δ bestimmen das Niveau der Semiinvarianten $s_{1,k}^\delta$ und haben folglich ebenfalls einen Einfluss auf die Iterationen, die der Regularisierungsalgorithmus benötigt. So wurden für alle Exponentialparameter bei der dritten Simulation zufallsbedingt wesentlich mehr Iterationen benötigt als bei der ersten Simulation.

Da der Exponentialparameter sichtbar die Anzahl der benötigten Iterationen und somit die Rechenzeit beeinflusst, sollte für weitere Rechnungen ein Exponentialparameter $\psi < 2$ gewählt werden. So haben sich Werte für ψ zwischen 0.1 und 1.5 als sehr

geeignet erwiesen. Ähnliche Resultate für den Exponentialparameter ψ wurden in [37] bei der Multiparameter-Regularisierung eines inversen Eigenwertproblems erzielt⁶.

Eine Kombination beider vorgestellten Modifikationen von Algorithmus 6.2.5 hat sich bei weiteren durchgeführten Rechnungen als besonders geeignet herausgestellt. So haben wir mit der modifizierten Abbruchbedingung $\|\underline{p}^{(j+1)} - \underline{p}^{(j)}\| \leq \varepsilon_2 = 0.000001$ und dem Exponentialparameter $\psi = 1.0$ eine deutliche Verringerung der benötigten Iterationen erreicht.

ψ	Iterationen	μ^δ	σ^δ	λ^δ	μ_Y^δ	σ^δ
0.1	281	0.222748	0.200587	10.671272	0.095210	0.195489
0.2	135	0.222748	0.200587	10.671272	0.095210	0.195489
0.3	106	0.222748	0.200587	10.671271	0.095210	0.195489
0.4	74	0.222749	0.200587	10.671274	0.095210	0.195489
0.5	89	0.222748	0.200587	10.671265	0.095210	0.195489
0.6	63	0.222748	0.200587	10.671270	0.095210	0.195489
0.7	76	0.222748	0.200587	10.671268	0.095210	0.195489
0.8	45	0.222741	0.200587	10.671110	0.095211	0.195489
0.9	49	0.222749	0.200587	10.671284	0.095210	0.195489
1	50	0.222748	0.200587	10.671271	0.095210	0.195489
1.1	45	0.222749	0.200587	10.671274	0.095210	0.195489
1.2	49	0.222748	0.200587	10.671280	0.095210	0.195489
1.3	62	0.222748	0.200587	10.671275	0.095210	0.195489
1.4	29	0.222745	0.200587	10.671100	0.095212	0.195489
1.5	46	0.222749	0.200587	10.671275	0.095210	0.195489
1.6	48	0.222748	0.200587	10.671275	0.095210	0.195489
1.7	161	0.222748	0.200587	10.671269	0.095210	0.195489
1.8	95	0.222748	0.200587	10.671265	0.095210	0.195489
1.9	93	0.222747	0.200587	10.671256	0.095210	0.195489
2	83	0.222747	0.200587	10.671248	0.095211	0.195489
2.5	112	0.222748	0.200587	10.671274	0.095210	0.195489
3.5	847	0.222748	0.200587	10.671273	0.095210	0.195489
4.5	3000	0.048101	0.200152	11.475861	0.078967	0.167304

Tabelle 6.6: Geschätzte Parametervektoren $\underline{p}_\alpha^\delta$ der ersten Simulation

⁶Aufgrund der in [37] verwendeten Struktur des Optimierungsproblems entsprechen allerdings die dort berechneten Exponentialparameter den reziproken Werten $\frac{1}{\psi}$.

ψ	1. Sim.	2. Sim.	3. Sim.	4. Sim.	5. Sim.	
0.1	281	325	1115	167	751	527.8
0.2	135	241	655	94	484	321.8
0.3	106	366	854	69	264	331.8
0.4	74	302	719	49	185	265.8
0.5	89	309	555	48	189	238
0.6	63	272	576	54	153	223.6
0.7	76	272	490	46	105	197.8
0.8	45	488	1013	42	142	346
0.9	49	292	1468	61	150	404
1	50	179	2533	69	91	584.4
1.1	45	551	559	36	137	265.6
1.2	49	648	786	52	137	334.4
1.3	62	947	1126	34	157	465.2
1.4	29	519	3095	66	111	764
1.5	46	439	808	62	181	307.2
1.6	48	564	1506	77	127	464.4
1.7	161	862	3067	50	165	861
1.8	95	433	5000	159	159	1169.2
1.9	93	997	3174	59	438	952.2
2	83	1233	5000	72	5000	2277.6
2.5	112	3000	3000	3000	354	1893.2
3.5	847	3000	3000	3000	3000	2569.4
4.5	3000	3000	3000	3000	3000	3000

Tabelle 6.7: Vergleich der benötigten Iterationen aller fünf Simulationen in Abhängigkeit des Exponentialparameters Ψ

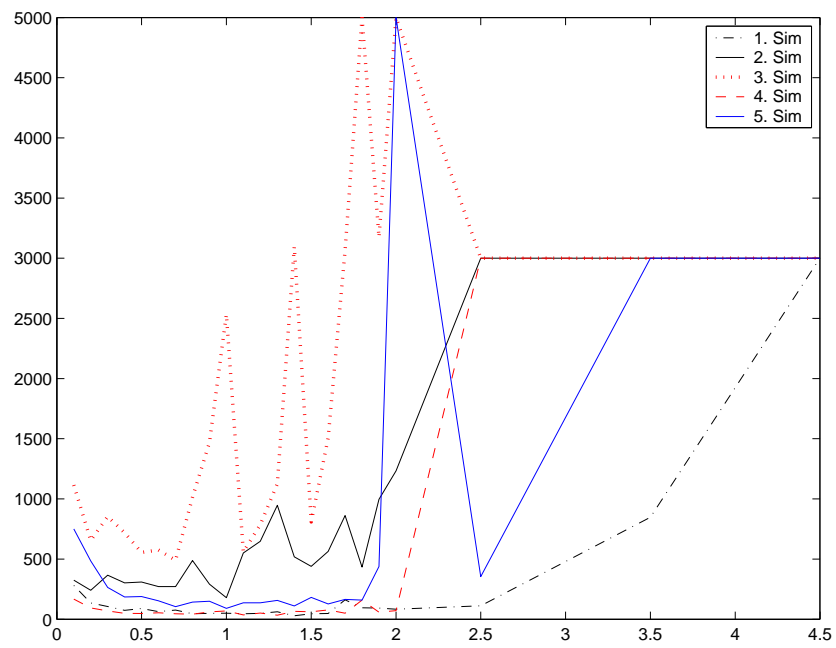


Abbildung 6.1: Benötigte Iterationen aller fünf Simulationen in Abhängigkeit des Exponentialparameters Ψ

6.4 Vergleichsrechnung zur Maximum-Likelihood-Methode

Zum Ende dieses Kapitels wollen wir noch einige Ergebnisse vergleichen, die wir einerseits mittels der Multiparameter-Regularisierung und andererseits durch die Maximum-Likelihood-Methode erhalten haben. Bisher haben wir die empirische Dichtefunktion bzw. den zugehörigen Datenvektor \underline{z}^δ durch Störung der exakten Dichtefunktion bzw. des Datenvektors \underline{z}^* erzeugt. In praktischen Situationen bestimmt man die empirische Dichtefunktion natürlich direkt aus den vorliegenden Aktienkursen bzw. den zugehörigen Returns. Daher simulieren wir in dieser abschließenden Fallstudie für einen Parametervektor \underline{p}^* eine Preistrajektorie und bestimmen daraus die Returns. Aus diesen Returns berechnen wir die empirischen Semiinvarianten $s_{1,1}^\delta, \dots, s_{1,l}^\delta$ sowie die empirische Dichtefunktion. Zur Bestimmung der empirischen Dichtefunktion wählen wir die Zerlegung $\tilde{x}_0 < \tilde{x}_1 < \dots < \tilde{x}_n$ aus Abschnitt 5.1 und bezeichnen die empirische Dichtefunktion daher im Folgenden mit f_n^δ .

Wir wollen die von uns mit der Multiparameter-Regularisierung erzielten Ergebnisse mit einer herkömmlichen statistischen Methode in Konkurrenz treten lassen. Daher verwenden wir zum Vergleich die Maximum-Likelihood-Methode, wobei wir nach Honoré [41] die Optimierung auf den Definitionsbereich (3.15) einschränken⁷. In dieser abschließenden Fallstudie haben wir 10000, 1000 bzw. 250 Returns für den Parametervektor $\underline{p}^* \in D_{\max}$ simuliert und aus diesen Daten mittels beider Verfahren den Parametervektor \underline{p}^δ geschätzt. Um einen zuverlässigen Vergleich garantieren zu können, wurde diese Prozedur mehrfach wiederholt.

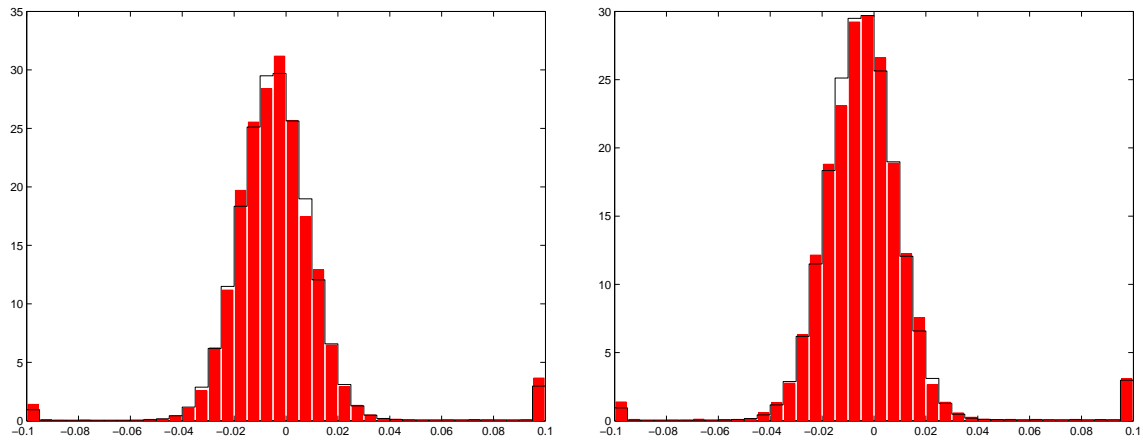


Abbildung 6.2: Empirische Dichtefunktionen f_n^δ (dargestellt als rotes Histogramm) aus je 10000 simulierten Tagesreturns im Vergleich zur diskretisierten exakten Dichtefunktion \tilde{f}_n^* (schwarze Linie)

Bei der ersten Simulationsstudie haben wir für $\underline{p}^* = (0.1, 0.2, 10, 0.1, 0.2)$ 10-mal jeweils 10000 Returns simuliert. In Abbildung 6.2 haben wir zwei der zehn empirischen

⁷Für die Multiparameter-Regularisierung haben wir wieder $D = D_{\max}$ gesetzt, während bei der Maximum-Likelihood-Methode eine Einschränkung des Definitionsbereichs numerisch notwendig war. Dort haben wir alle Parameterwerte betragsmäßig durch 1000 beschränkt.

Dichtefunktionen, die sich aus den simulierten Returns ergeben haben, als Histogramm dargestellt. Um den Datenfehler δ mit $\|f_n^\delta - f_n^*\| \leq \delta$ zu verdeutlichen, wurde die zu \underline{p}^* gehörende exakte diskretisierte Dichtefunktion \tilde{f}_n^* dieser Grafik hinzugefügt. Trotz der hohen Anzahl verwendeter Daten ist eine leichte Störung der Dichtefunktion zu erkennen.

Die Ergebnisse aller 10 Simulationen sind in Abbildung 6.3 dargestellt. Die linke Seite zeigt dabei die Ergebnisse aller fünf Werte des geschätzten Parametervektors der Maximum-Likelihood-Schätzung und die rechte Seite die Ergebnisse der Multiparameter-Regularisierung. So enthält beispielsweise der jeweils erste Balken die Ergebnisse des Diffusionsparameters μ . Dabei symbolisieren die blauen Striche jeweils einen geschätzten Wert des entsprechenden Parameters. Die rote gestrichelte Linie stellt den Median aller geschätzten Werte des jeweiligen Parameters dar, und die rote durchgezogene Linie repräsentiert den exakten Parameterwert. Beim Vergleich der beiden Teilgrafiken sollten die jeweiligen Skalen beachtet werden, da sich diese mitunter stark unterscheiden.

Abbildung 6.3 zeigt, dass sowohl die Maximum-Likelihood-Methode als auch die Multiparameter-Regularisierung bei der Verwendung von 10000 Daten im Mittel sehr gute Schätzungen liefern. Die einzelnen geschätzten Werte schwanken bei beiden Verfahren in der gleichen Größenordnung um die exakten Werte. Folglich unterscheiden sich bei dieser ersten Simulationsstudie kaum die Skalen, auf denen die Ergebnisse der Maximum-Likelihood-Schätzung bzw. der Multiparameter-Regularisierung abgetragen sind.

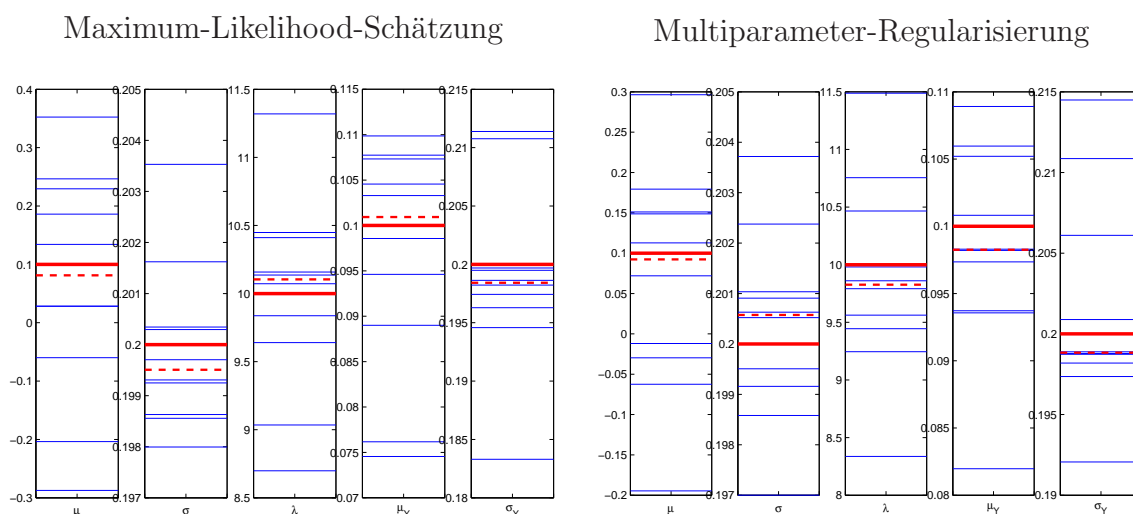


Abbildung 6.3: Ergebnisse aller 10 Simulationen von jeweils 10000 Tageskursen

In den meisten praktischen Situationen stehen natürlich keine 10000 Datensätze zur Verfügung, so dass wir im Weiteren weniger Datensätze simulieren werden.

Bei der zweiten Studie haben wir insgesamt 100-mal jeweils 1000 Returns für den Parametervektor $\underline{p}^* = (0.1, 0.2, 10, 0.1, 0.2)$ simuliert. Abbildung 6.4 zeigt zwei empirische Dichtefunktionen der jeweils 1000 simulierten Returns. Im Vergleich zur Simulation von 10000 Returns ist ein deutlich größerer Datenfehler δ zu erkennen. Ferner zeigt ein Vergleich der linken und rechten Grafik, dass die Abweichungen zwischen der empiri-

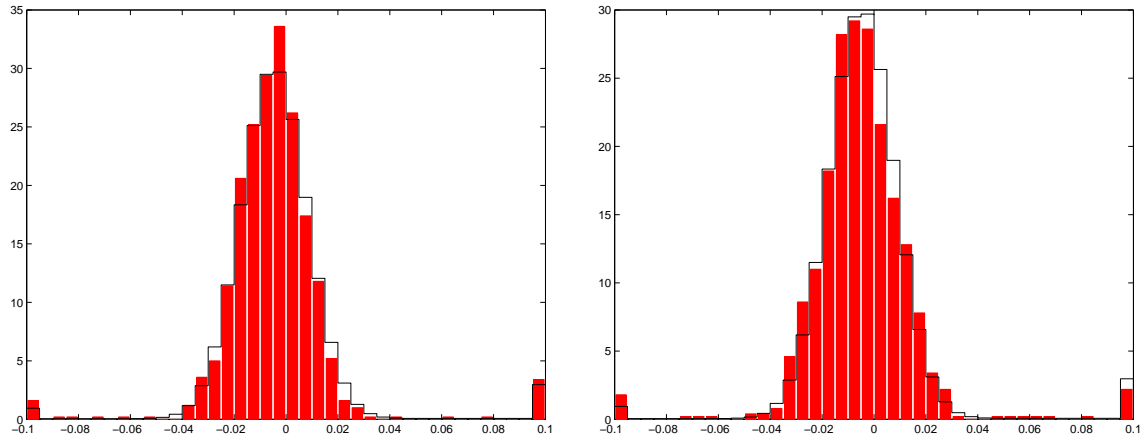


Abbildung 6.4: Empirische Dichtefunktionen aus je 1000 simulierten Tagesreturns

schen Dichtefunktion f_n^δ und der exakten diskretisierten Dichtefunktion \tilde{f}_n^* mit jeder Simulation eine neue Charakteristik aufweisen. Dieser größere Datenfehler bewirkt bei beiden Verfahren eine höhere Schwankungsbreite der geschätzten Parameter, wie in Abbildung 6.5 zu erkennen ist. Allerdings reagiert die Maximum-Likelihood-Schätzung deutlich sensibler auf Datenfehler als die Multiparameter-Regularisierung. So wurde der Diffusionsparameter σ mit dem Maximum-Likelihood-Verfahren teilweise mit $\sigma^\delta \approx 0.6$ deutlich zu hoch geschätzt. Bei einigen Simulationen ist die Sprungintensität auf sehr kleine Werte geschätzt worden. Falls $\lambda \approx 0$ gilt, so haben die Werte der Sprungparameter μ_Y und σ_Y keine Bedeutung. Es wurden aber bei der Maximum-Likelihood-Schätzung deutlich mehr Werte mit $\mu_Y^\delta < -20$ und $\sigma_Y^\delta > 0.6$ falsch geschätzt, als die Fehlschätzungen $\lambda^\delta \approx 0$ rechtfertigen würden, was auf Instabilitätsprobleme bei diesem Schätzverfahren schließen lässt.

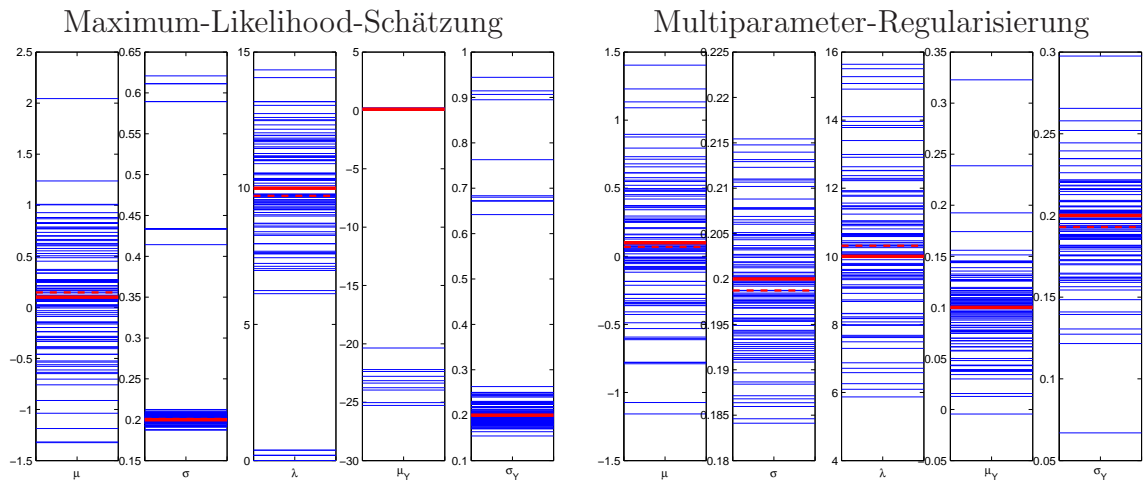


Abbildung 6.5: Ergebnisse aller 100 Simulationen von jeweils 1000 Tageskursen

Natürlich schwanken auch die Schätzungen der Multiparameter-Regularisierung um die exakten Werte. Jedoch ist deutlich zu erkennen, dass wir mit der Multiparameter-Regularisierung im Vergleich zur Maximum-Likelihood-Schätzung eine Stabilisierung erreichen konnten. Die geringere Schwankungsbreite der mittels Multiparameter-Re-

gularisierung geschätzten Parameter ist deutlich an den verschiedenen Skalierungen beider Grafiken zu erkennen. Während der Parameter σ bei der Maximum-Likelihood-Methode Werte zwischen 0.15 und 0.65 angenommen hat, variierten die 100 Schätzungen der Multiparameter-Regularisierung zwischen 0.18 und 0.225. Noch deutlichere Unterschiede der Skalen bzw. Schwankungsbreiten zeigen sich bei den Sprungparametern μ_Y und σ_Y . Die mittlere Sprunghöhe μ_Y wurde bei der Multiparameter-Regularisierung zwischen -0.05 und 0.35 geschätzt, während die Werte bei der Maximum-Likelihood-Schätzung zwischen -30 und 5 lagen. Die Skala der Standardabweichung der Sprunghöhen σ_Y liegt für die Maximum-Likelihood-Schätzung zwischen 0.1 und 1 und die der Multiparameter-Regularisierung zwischen 0.05 und 0.3 .

Bei dem nächsten Beispiel wollen wir nun die realistische Situation nachstellen, dass wir Tageskurse von einem Handelsjahr als Grundlage zur Schätzung der Parameter zur Verfügung haben. Daher simulieren wir für den Parametervektor $\underline{p}^* = (0.1, 0.2, 10, 0.1, 0.2)$ nur 250 Returns und betrachten die Ergebnisse beider Verfahren. In Abbildung 6.6 haben wir nochmals zwei empirische Dichtefunktionen von jeweils 250 simulierten Returns dargestellt.

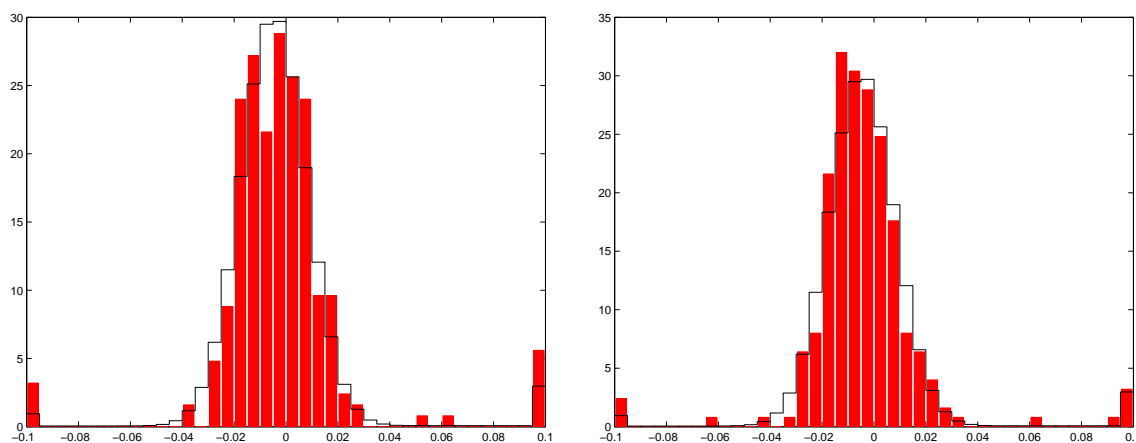


Abbildung 6.6: Empirische Dichtefunktionen aus je 250 simulierten Tagesreturns

Obige Abbildung zeigt deutlich, dass bei 250 Returns die empirische Dichtefunktion f_n^δ starke Abweichungen zur exakten Funktion \tilde{f}_n^* aufweist. Da bei der Maximum-Likelihood-Methode bereits bei 1000 Returns Stabilitätsprobleme auftreten, ist bei diesem größeren Datenfehler ebenfalls mit Instabilitätsphänomenen zu rechnen. Dies motiviert in praktischen Situationen die Verwendung eines Regularisierungszugangs des inversen Problems der Parameterbestimmung in Jump-Diffusions Prozessen. Die Abbildungen 6.7 und 6.8 zeigen die Ergebnisse von jeweils 20 bzw. 100 Simulationen. Prinzipiell ergibt sich das gleiche Bild, wie wir es bereits bei der zweiten Studie mit 1000 simulierten Returns gesehen haben.

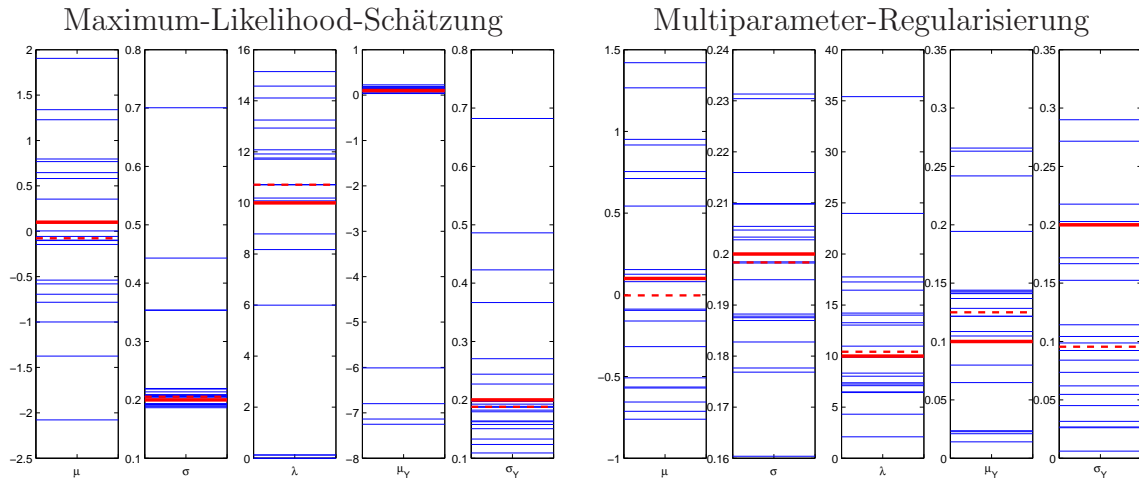


Abbildung 6.7: Ergebnisse aller 20 Simulationen von jeweils 250 Tageskursen

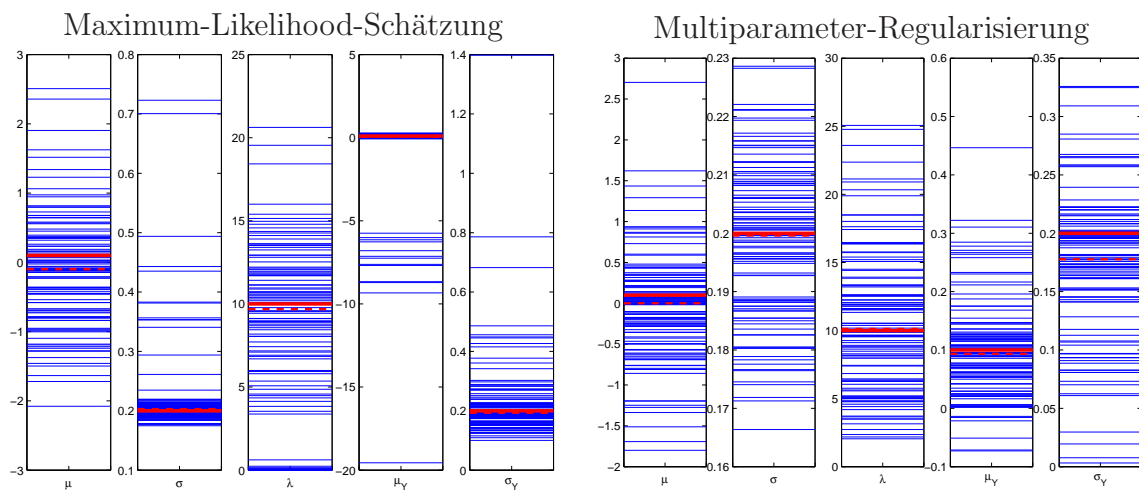


Abbildung 6.8: Ergebnisse aller 100 Simulationen von jeweils 250 Tageskursen

Zum Abschluss präsentieren wir noch die Ergebnisse einer weiteren Studie mit einem anderen Parametervektor $\underline{p}^* = (0.2, 0.2, 50, 0.05, 0.05)$. Hierfür haben wir 50-mal jeweils 250 Returns simuliert. Die Ergebnisse, die in Abbildung 6.9 dargestellt sind, zeigen ähnliche Effekte, wie wir sie bei dem von uns betrachteten Standardbeispiel $\underline{p}^* = (0.1, 0.2, 10, 0.1, 0.2)$ beobachtet haben.

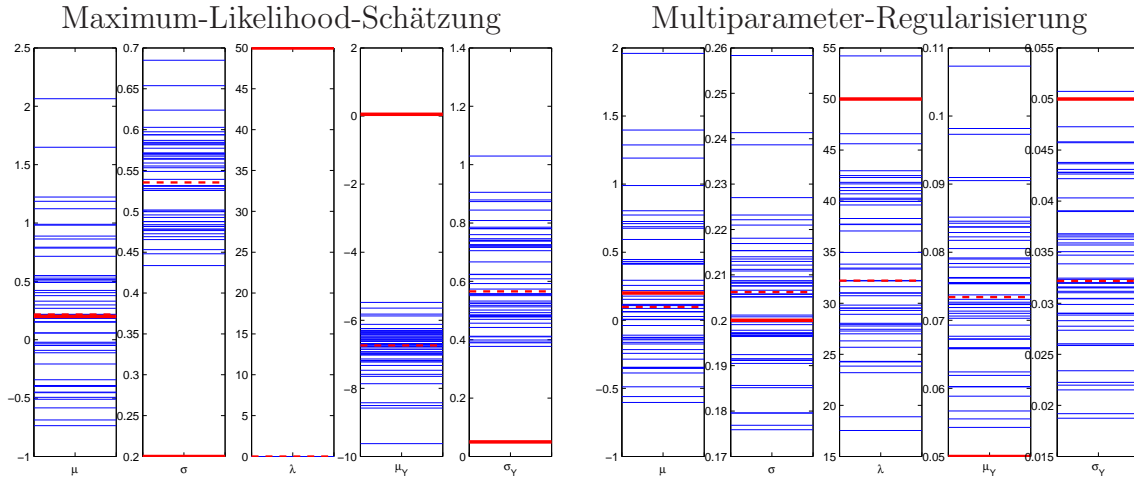


Abbildung 6.9: Ergebnisse aller 50 Simulationen von jeweils 250 Tageskursen mit $\underline{p}^* = (0.2, 0.2, 50, 0.05, 0.05)$

Die Ergebnisse der Multiparameter-Regularisierung zeigen mehr Stabilität als die der Maximum-Likelihood-Schätzung. Bei diesem letzten Beispiel ist bei der Maximum-Likelihood-Schätzung das Phänomen der asymptotischen Nichtinjektivität zu beobachten, welches wir in Abschnitt 4.3 beschrieben haben. Die vielen kleineren Sprünge werden hier fälschlicherweise als hohe Volatilität des Diffusionsteils interpretiert. Die mit $\lambda^\delta \approx 0$ geschätzte Sprungintensität weicht deutlich von dem exakten Wert $\lambda^* = 50$ ab. Daher sind die Schätzwerte der restlichen Sprungparameter μ_Y und σ_Y bedeutungslos. Die bei der Maximum-Likelihood-Methode deutlich zu hoch geschätzten Werte σ^δ des Parameters σ machen deutlich, dass der Volatilitätsanteil der nicht identifizierten Sprungkomponente dem Diffusionsteil zugeordnet wurde. Bei diesem schwierigen Beispiel zeigt auch die Multiparameter-Regularisierung Schwächen. So liegt der Stichprobenmedian der Sprungintensität bei 31, was immer noch eine deutliche Abweichung zu $\lambda^* = 50$ darstellt. Trotz alledem wird der Sprungteil im Vergleich zur Maximum-Likelihood-Methode relativ gut eingeschätzt, was auch in den Schätzwerten des Parameters σ zu sehen ist.

Kapitel 7

Anwendungen in der Overreaction-Theorie

Bei vielen finanzmathematischen Fragestellungen, beispielsweise der Bestimmung von Optionspreisen in einem konkreten Kursmodell, müssen die Parameter dieses Kursmodells an die Marktdaten angepasst werden. Zum Abschluss dieser Arbeit wollen wir eine spezielle Anwendung der Parameterschätzung für Jump-Diffusions-Prozesse vorstellen.

Bei der Betrachtung verschiedener Finanzmärkte kann man teilweise irrationales Verhalten der Marktteilnehmer beobachten, wenn völlig unerwartete Neuigkeiten bekannt werden. Die Kurse der betroffenen Finanztitel steigen bzw. fallen in solchen Situationen unerklärlich stark. Dieses Phänomen der Überreaktion (Overreaction-Effekt) der Marktteilnehmer, das sowohl bei einzelnen Aktien als auch bei Aktienindizes zu beobachten ist, wurde erstmals in den 80er-Jahren von De Bondt und Thaler in [14] diskutiert. In dieser Arbeit wurden die folgenden zwei Thesen, die Richtung und Ausmaß der Überreaktion beschreiben, aufgestellt:

1. Extremen Bewegungen in den Kursen, die durch Überreaktionen der Marktteilnehmer entstehen, folgen Kursbewegungen in die entgegengesetzte Richtung.
2. Je extremer die Kursbewegung, desto extremer fällt die Gegenbewegung aus.

Seither sind zahlreiche Arbeiten (u.a. [9], [27], [29], [31], [49], [57], [58]) erschienen, die diese Thesen für bestimmte Finanzmärkte näher untersuchen und spezifizieren. In [9] wurden, in Abhängigkeit von der Dauer der Extrembewegung sowie der Gegenbewegung, drei verschiedene Kategorien des Overreaction-Effektes eingeführt. Wir werden uns im Weiteren auf den kurzzeitigen Overreaction-Effekt, der ein paar Tage andauert, konzentrieren.

Die Identifikation extremer Kursbewegungen kann man auch als inverses Diagnoseproblem interpretieren. Wir stellen im Folgenden einen Indikator vor, der solche Überreaktionen anzeigen soll (siehe auch [39] und [18]). Dieser Overreaction-Indikator basiert auf der Berechnung von Wahrscheinlichkeiten gewisser Kursbewegungen. Dabei sind diese Wahrscheinlichkeiten vom betrachteten Kursmodell und den zugehörigen Parametern, die dieses Modell näher quantifizieren, abhängig. Um die Wahrscheinlichkeiten

von Kursbewegungen genau zu berechnen, benötigen wir daher eine zuverlässige und stabile Schätzung der Parameter aus den vorliegenden Marktdaten.

7.1 Ein Overreaction-Indikator

Wir betrachten die Marktdaten $(S_{t_0}^\delta, S_{t_1}^\delta, \dots, S_{t_n}^\delta)$ und die zugehörigen Returns $r_{t_i, t_i+\tau}^\delta = \ln \frac{S_{t_i+\tau}^\delta}{S_{t_i}^\delta}$. Da bei einer Overreaction-Situation die Kurse überdurchschnittlich stark steigen bzw. fallen, haben die realisierten Returns sehr große bzw. kleine Werte. Die Wahrscheinlichkeit, dass die Returns $r_{t_i, t_i+\tau}$ mit Lag τ solche großen bzw. kleinen Werte $r_{t_i, t_i+\tau}^\delta$ annehmen, sollte daher sehr klein sein. Deswegen ordnen wir jedem realisierten Return $r_{t_i, t_i+\tau}^\delta$ die Wahrscheinlichkeit $\hat{\mathbb{P}}_{t_i, t_i+\tau}$ mit

$$\hat{\mathbb{P}}_{t_i, t_i+\tau} := \begin{cases} \mathbb{P}(r_{t_i, t_i+\tau} > r_{t_i, t_i+\tau}^\delta) & \text{falls } r_{t_i, t_i+\tau}^\delta > 0 \\ \mathbb{P}(r_{t_i, t_i+\tau} < r_{t_i, t_i+\tau}^\delta) & \text{falls } r_{t_i, t_i+\tau}^\delta \leq 0 \end{cases} \quad (7.1)$$

zu. Wir nehmen an, dass $(S_{t_0}^\delta, S_{t_1}^\delta, \dots, S_{t_n}^\delta)$ die Realisierung eines Jump-Diffusions-Prozesses ist, der durch den Parametervektor $\underline{p} \in D_{\max}$ näher spezifiziert wird. Wegen (2.14) gilt dann

$$\hat{\mathbb{P}}_{t_i, t_i+\tau} = \begin{cases} 1 - F_{r_\tau}(r_{t_i, t_i+\tau}^\delta, \underline{p}) = 1 - \sum_{j=0}^{\infty} \frac{e^{(-\lambda\tau)}(\lambda\tau)^j}{j!} \Phi\left(\frac{r_{t_i, t_i+\tau}^\delta - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right), & r_{t_i, t_i+\tau}^\delta > 0 \\ F_{r_\tau}(r_{t_i, t_i+\tau}^\delta, \underline{p}) = \sum_{j=0}^{\infty} \frac{e^{(-\lambda\tau)}(\lambda\tau)^j}{j!} \Phi\left(\frac{r_{t_i, t_i+\tau}^\delta - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right), & r_{t_i, t_i+\tau}^\delta \leq 0. \end{cases} \quad (7.2)$$

Zur praktischen Berechnung der Wahrscheinlichkeit (7.2) müssen wir die Summation der Verteilungsfunktion $F_{r_\tau}(\cdot, \underline{p}) = \mathbb{P}(r_\tau \leq x)$ nach $N+1$ Schritten analog zur Dichtefunktion (3.12) abbrechen, so dass wir $F_{r_\tau}(\cdot, \underline{p})$ durch die Approximation

$$F_{r_\tau}^N(x, \underline{p}) := \sum_{j=0}^N \frac{e^{(-\lambda\tau)}(\lambda\tau)^j}{j!} \Phi\left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}}\right)$$

ersetzen. In Analogie zu Lemma 3.2.2 können wir den Approximationsfehler abschätzen.

Lemma 7.1.1 *Der Approximationsfehler $E^N(x, \underline{p}) := |F_{r_\tau}^N(x, \underline{p}) - F_{r_\tau}(x, \underline{p})|$ lässt sich für alle $x \in \mathbb{R}$ durch*

$$E^N(x, \underline{p}) \leq 1 - e^{-\lambda\tau} \sum_{j=0}^N \frac{(\lambda\tau)^j}{j!}$$

abschätzen.

Beweis. Die Behauptung folgt unmittelbar aus der Beziehung

$$\begin{aligned} E^N(x, \underline{p}) &= \sum_{j=N+1}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} \Phi \left(\frac{x - (\tilde{\mu}\tau + j\mu_Y)}{\sqrt{\sigma^2\tau + j\sigma_Y^2}} \right) \\ &\leq \sum_{j=N+1}^{\infty} e^{-\lambda\tau} \frac{(\lambda\tau)^j}{j!} = 1 - e^{-\lambda\tau} \sum_{j=0}^N \frac{(\lambda\tau)^j}{j!} \end{aligned}$$

□

Bemerkung 7.1.2 Für den Spezialfall $\underline{p} \in D \setminus \hat{D}$ vereinfacht sich das Kursmodell der Jump-Diffusions-Prozesse zu dem Modell der geometrischen Brownschen Bewegung, bei dem die Returns $r_{t_i, t_i+\tau}$ normalverteilt sind. Folglich reduziert sich (7.2) wegen (4.21) zu

$$\hat{\mathbb{P}}_{t_i, t_i+\tau} = \begin{cases} 1 - \Phi \left(\frac{r_{t_i, t_i+\tau}^\delta - \tilde{\mu}\tau}{\sqrt{\sigma^2\tau}} \right) & \text{falls } r_{t_i, t_i+\tau}^\delta > 0 \\ \Phi \left(\frac{r_{t_i, t_i+\tau}^\delta - \tilde{\mu}\tau}{\sqrt{\sigma^2\tau}} \right) & \text{falls } r_{t_i, t_i+\tau}^\delta \leq 0 \end{cases}$$

und für jedes $N \geq 0$ gilt $E^N(x, \underline{p}) = 0$.

Da in praktischen Situationen der Parametervektor \underline{p} unbekannt ist, ersetzen wir ihn in (7.2) mit dem aus den Kursdaten $(S_{t_0}^\delta, S_{t_1}^\delta, \dots, S_{t_n}^\delta)$ geschätzten Parametervektor $\underline{p}^\delta = (\mu^\delta, \sigma^\delta, \lambda^\delta, \mu_Y^\delta, \sigma_Y^\delta) \in D_{\max}$. Dabei nutzen wir zur Berechnung des Parametervektors \underline{p}^δ die in Kapitel 6 vorgestellte Multiparameter-Regularisierung. Mit der Wahrscheinlichkeit (7.1) führen wir nun den folgenden Overreaction-Indikator ein.

Definition 7.1.3 Die reale Kursbewegung vom Zeitpunkt t_{i-3} zum Zeitpunkt t_i nennen wir qualifizierte 3-Tages Bewegung, falls die folgenden vier Bedingungen erfüllt sind:

1. Die Kurse $S_{t_{i-3}}^\delta$, $S_{t_{i-2}}^\delta$, $S_{t_{i-1}}^\delta$ und $S_{t_i}^\delta$ bilden eine streng monoton fallende oder steigende Folge.
2. Die absolute Rendite beträgt mindestens β , d.h. $\frac{|S_{t_i}^\delta - S_{t_{i-3}}^\delta|}{S_{t_{i-3}}^\delta} \geq \beta$.
3. Eine der Wahrscheinlichkeiten $\hat{\mathbb{P}}_{t_{i-3}, t_{i-2}}$ und $\hat{\mathbb{P}}_{t_{i-2}, t_{i-1}}$ ist kleiner als ϵ_1 .
4. Die Wahrscheinlichkeit $\hat{\mathbb{P}}_{t_{i-3}, t_i}$ ist kleiner als ϵ_2 .

Wir nehmen an, dass eine Overreaction-Situation vorliegt, wenn eine solche qualifizierte 3-Tages Bewegung festgestellt ist.

Abbildung 7.1 zeigt zwei Overreaction-Indikationen bei dem DAX-Future, der im März 1999 fällig war¹. Zur Identifikation der qualifizierten 3-Tages Bewegung wurden die Einstellungen $\beta = 0.04$, $\epsilon_1 = 0.1$ und $\epsilon_2 = 0.2$ verwendet. Allerdings haben wir bei diesem

¹Da ein DAX-Future nur in den letzten drei Handelsmonaten liquide ist, haben wir nur in diesem Zeitraum Overreaction-Situationen analysiert.

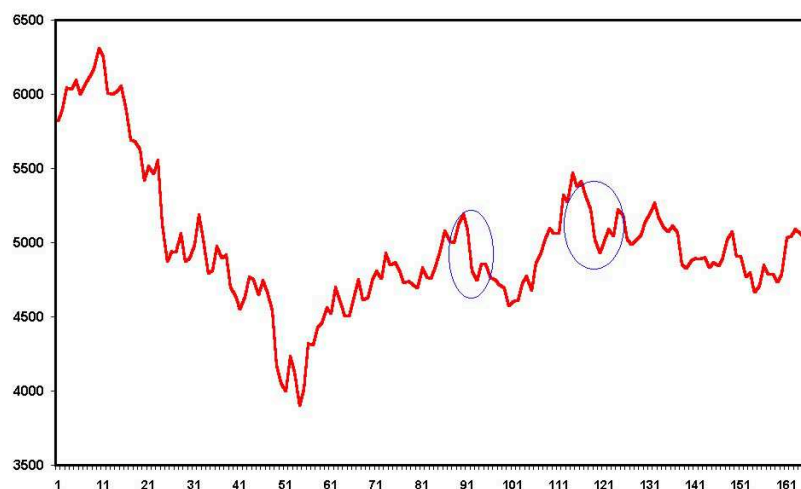


Abbildung 7.1: Qualifizierte 3-Tages Bewegungen beim DAX-Future 99/03

Beispiel den Spezialfall der geometrischen Brownschen Bewegung explizit vorausgesetzt. In beiden Fällen sind unterschiedlich ausgeprägte Gegenbewegungen zu erkennen.

Das frühzeitige Erkennen von Overreaction-Situationen kann in verschiedene Handelsstrategien umgesetzt werden (siehe auch [20]), indem die Information der erwarteten Gegenbewegung ausgenutzt wird. Falls die Kurse stark gefallen sind, erwartet man einen deutlichen Kursanstieg, so dass man in diesem Fall ein Long-Investment eingeht indem das Finanzgut selbst oder ein Derivat, wie z.B. eine Option, ein Future oder ein Zertifikat gekauft wird. Andernfalls erwartet man nach einem starken Kursanstieg fallende Kurse und geht ein Short-Investment ein, d.h. verkauft das Finanzgut oder handelt ein passendes Derivat. Allerdings scheint es schwierig solche Handelsstrategien als seriösen Bewertungsmaßstab für die erwartete Gegenbewegung zu verwenden. Einerseits ist der Gewinn des Investments abhängig vom Handelseinstieg, d.h. dem Zeitpunkt der Indikation einer Overreaction-Situation und andererseits wird der Gewinn auch maßgeblich vom Zeitpunkt des Ausstiegs aus dieser Position geprägt. Nach jeder Overreaction-Situation, die durch irrationales Verhalten der Marktteilnehmer ausgelöst wird, ist das Marktgeschehen nach einer gewissen Erholungsphase wieder durch rationales Verhalten der Händler geprägt. Da die Dauer einer Gegenbewegung nicht vorhersagbar ist, kann beispielsweise eine klassische Overreaction-Situation durch zu langes Halten der Investment-Position mit einem Verlust geschlossen werden. Durch geschicktes Handeln kann andererseits auch bei jedem beliebigen Kursverlauf ein Gewinn erzielt werden, so dass eine klare Zuordnung des erzielten Gewinnes auf die Indikation sehr schwierig scheint. Daher werden wir im Weiteren andere Maße zur Analyse der Gegenbewegung betrachten.

7.2 Maße für die Gegenbewegung

Der Overreaction-Effekt ist nicht nur für Handelsstrategien interessant. Die Identifikation von Overreaction-Situationen ist auch aus wissenschaftlicher Sicht bedeutsam, da ein seriöser Nachweis dieses Effektes helfen kann, das Marktgeschehen genauer zu erklären. Wenn die Hypothesen der Overreaction-Theorie gelten, so muss nach einer extremen Bewegung der Kurse, die wir durch eine qualifizierte 3-Tages Bewegung identifizieren, eine entsprechende Gegenbewegung folgen. Daher beobachten wir nach jeder qualifizierten 3-Tages Bewegung, die zum Zeitpunkt t_i endet, die Entwicklung der darauffolgenden Kurse des Assets. Falls der Overreaction-Indikator eine Abwärtsbewegung angezeigt hat, erwarten wir steigende Kurse, so dass wir für jedes Lag $\tau > 0$ die Returns $r_{t_i, t_i+\tau}^\delta$ nach dem Indikationszeitpunkt t_i betrachten. Hat dagegen der Overreaction-Indikator zum Zeitpunkt t_i eine Aufwärtsbewegung angezeigt, betrachten wir die negativen kumulierten Returns, so dass auch diese Kenngröße hohe Werte annimmt, sobald die Kurse fallen. In Analogie zu [29] bezeichnen wir dieses Maß für die Gegenbewegung einer qualifizierten 3-Tages Bewegung $S_{t_{i-3}}^\delta$, $S_{t_{i-2}}^\delta$, $S_{t_{i-1}}^\delta$ und $S_{t_i}^\delta$ mit CAR (cumulative abnormal return)

$$CAR(t_i, \tau) := \begin{cases} r_{t_i, t_i+\tau}^\delta = \ln \frac{S_{t_i+\tau}^\delta}{S_{t_i}^\delta} & \text{falls } S_{t_{i-3}}^\delta > S_{t_i}^\delta \\ -r_{t_i, t_i+\tau}^\delta = \ln \frac{S_{t_i}^\delta}{S_{t_i+\tau}^\delta} & \text{falls } S_{t_{i-3}}^\delta < S_{t_i}^\delta. \end{cases}$$

Wir werden uns jetzt im Weiteren auf den Fall einer qualifizierten Abwärtsbewegung konzentrieren, da sich alle betrachteten Kenngrößen einer Aufwärtsbewegung analog berechnen lassen.

Definition 7.2.1 (ACAR) *Es seien $t_{ind_1}, t_{ind_2}, \dots, t_{ind_M}$ die Zeitpunkte aller Indikationen qualifizierter Abwärtsbewegungen. Wir bezeichnen das arithmetische Mittel aller kumulierten Returns mit ACAR (averaged cumulative abnormal return)*

$$ACAR(\tau) := \frac{1}{M} \sum_{i=1}^M CAR(t_{ind_i}, \tau). \quad (7.3)$$

Die durchschnittlich kumulierten Returns ACAR sind vergleichbar mit dem Erwartungswert der (stationären) Returns eines Jump-Diffusions-Prozesses

$$ECAR(\tau) := \mathbb{E}r_\tau = (\tilde{\mu} + \lambda\mu_Y) \tau.$$

In der Literatur zur Overreaction-Theorie wird die Entwicklung des Maßes $ACAR(\tau)$ für die folgenden Handelstage nach extremen Kursbewegungen analysiert, um eine Gegenbewegung nachzuweisen. Durch den Vergleich des Maßes ACAR mit dem Erwartungswert $ECAR(\tau)$ können wir die Gegenbewegung viel kritischer bewerten, da die Werte $ECAR(\tau)$ die durchschnittliche Entwicklung der kumulierten Returns unabhängig von Overreaction-Situationen angeben.

Ferner führen wir, in Analogie zur mittleren Entwicklung der kumulierten Returns, zusätzlich ein Maß ein, welches die bestmögliche Entwicklung der kumulierten Returns beschreibt (siehe [18]).

Definition 7.2.2 (ACBR) Wir definieren die bestmöglichen kumulierten Returns *CBR* (cumulative best case return) einer qualifizierten Abwärtsbewegung durch

$$CBR(t_i, \tau) := \max_{0 \leq t \leq \tau} r_{t_i, t_i+t}^\delta = \max_{0 \leq t \leq \tau} \ln \frac{S_{t_i+t}^\delta}{S_{t_i}^\delta}$$

und bezeichnen das Mittel aller bestmöglichen Returns nach den Indikationen qualifizierter 3-Tages Abwärtsbewegungen zu den Zeitpunkten $t_{ind_1}, t_{ind_2}, \dots, t_{ind_M}$ mit *ACBR* (averaged cumulative best case return)

$$ACBR(\tau) := \frac{1}{M} \sum_{i=1}^M CBR(t_{ind_i}, \tau). \quad (7.4)$$

Das Maß ACBR ist mit dem Erwartungswert des Laufzeit-Maximums (Running-Maximum) der Returns

$$ECBR(\tau) := \mathbb{E} \sup_{0 \leq t \leq \tau} r_t \quad (7.5)$$

vergleichbar.

Neben der bestmöglichen Entwicklung beobachten wir ferner den Return, der im schlechtesten Fall realisiert wird.

Definition 7.2.3 (ACWR) Wir definieren analog zu den bestmöglichen kumulierten Returns die Kenngröße *CWR* (cumulative worst case return) einer qualifizierten Abwärtsbewegung durch

$$CWR(t_i, \tau) := \min_{0 \leq t \leq \tau} r_{t_i, t_i+t}^\delta = \min_{0 \leq t \leq \tau} \ln \frac{S_{t_i+t}^\delta}{S_{t_i}^\delta}$$

und das Mittel *ACWR* (averaged cumulative worst case return) mit

$$ACWR(\tau) := \frac{1}{M} \sum_{i=1}^M CWR(t_{ind_i}, \tau). \quad (7.6)$$

In Analogie zum bestmöglichen Fall vergleichen wir das Maß ACWR mit dem Erwartungswert des Laufzeit-Minimums (Running-Minimum) der Returns

$$ECWR(\tau) := \mathbb{E} \inf_{0 \leq t \leq \tau} r_t. \quad (7.7)$$

Mit den Maßen ACAR, ACBR und ACWR kann nun die Hypothese, dass extremen Kursbewegungen entsprechende Gegenbewegungen folgen, überprüft werden. Ferner können wir durch Vergleich der Werte ACAR und ECAR, ACBR und ECBR sowie ACWR und ECWR feststellen, ob die tatsächlichen Gegenbewegungen überdurchschnittlich ausfallen. Daher werden wir im nächsten Abschnitt die Möglichkeiten zur Berechnung der Erwartungswerte der Laufzeitextrema (7.5) und (7.7) diskutieren.

7.3 Erwartungswerte der Laufzeitextrema

Die explizite Berechnung der Erwartungswerte der Laufzeitextrema ist ein schwieriges mathematisches Problem. Wir stellen hier den Zugang aus [68] zur Berechnung der Verteilungsfunktionen des Running-Maximums bzw. Running-Minimums vor, mit deren Kenntnis die Erwartungswerte berechnet werden können. Hierfür definieren wir zunächst

$$q_+(\theta, c) := \int_0^\infty e^{-\theta\tau} \mathbb{P}(\sup_{0 \leq t \leq \tau} r_t < c) d\tau \quad (7.8a)$$

$$q_-(\theta, c) := \int_0^\infty e^{-\theta\tau} \mathbb{P}(\inf_{0 \leq t \leq \tau} r_t < c) d\tau \quad (7.8b)$$

und

$$\tilde{q}_+(\theta, z) := \int_{-\infty}^\infty e^{ixz} dq_+(\theta, x) \quad (7.9a)$$

$$\tilde{q}_-(\theta, z) := \int_{-\infty}^\infty e^{ixz} dq_-(\theta, x). \quad (7.9b)$$

Die Funktionen $q_+(\theta, c)$ und $q_-(\theta, c)$ sind die bezüglich τ Laplace-Transformierten der gesuchten Verteilungsfunktionen $\mathbb{P}(\sup_{0 \leq t \leq \tau} r_t < c)$ bzw. $\mathbb{P}(\inf_{0 \leq t \leq \tau} r_t < c)$. Ferner entsprechen die Funktionen $\tilde{q}_+(\theta, c)$ und $\tilde{q}_-(\theta, c)$ den Fourier-Transformationen von $q_+(\theta, c)$ und $q_-(\theta, c)$. Der folgende Satz aus [68, Theorem 28° (p. 210)] zeigt den Zusammenhang zwischen den Funktionen $\tilde{q}_+(\theta, c)$ und $\tilde{q}_-(\theta, c)$ und der Verteilungsfunktion $F_{r_\tau}(\cdot, \underline{p})$ bzw. Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ der Returns.

Satz 7.3.1 *Die Funktionen $\tilde{q}_+(\theta, c)$ und $\tilde{q}_-(\theta, c)$ lassen sich aus der Verteilungs- bzw. Dichtefunktion der Returns eines Jump-Diffusions-Prozesses mit Parametervektor $\underline{p} \in D$ durch*

$$\begin{aligned} \tilde{q}_+(\theta, c) &= \frac{1}{\theta} \exp \left(\int_0^\infty \frac{1}{\tau} e^{-\theta\tau} \int_0^\infty (e^{icx} - 1) d_x F_{r_\tau}(x, \underline{p}) d\tau \right) \\ &= \frac{1}{\theta} \exp \left(\int_0^\infty \frac{1}{\tau} e^{-\theta\tau} \int_0^\infty (e^{icx} - 1) f_{r_\tau}(x, \underline{p}) dx d\tau \right) \end{aligned} \quad (7.10)$$

sowie

$$\begin{aligned} \tilde{q}_-(\theta, c) &= \frac{1}{\theta} \exp \left(\int_0^\infty \frac{1}{\tau} e^{-\theta\tau} \int_{-\infty}^0 (e^{icx} - 1) d_x F_{r_\tau}(x, \underline{p}) d\tau \right) \\ &= \frac{1}{\theta} \exp \left(\int_0^\infty \frac{1}{\tau} e^{-\theta\tau} \int_{-\infty}^0 (e^{icx} - 1) f_{r_\tau}(x, \underline{p}) dx d\tau \right). \end{aligned} \quad (7.11)$$

berechnen.

Mit den Formeln (7.10) und (7.11) können wir durch numerische Berechnung der jeweiligen inversen Fourier-Transformierten von (7.9) die Funktionen $q_+(\theta, z)$ bzw. $q_-(\theta, z)$ approximieren. Mittels inverser Laplace-Transformation von (7.8) können wir dann die gesuchten Verteilungsfunktionen $\mathbb{P}(\sup_{0 \leq t \leq \tau} r_t < c)$ bzw. $\mathbb{P}(\inf_{0 \leq t \leq \tau} r_t < c)$ numerisch berechnen. Allerdings zählt die inverse Laplace-Transformation zu den inkorrekten inversen Problemen (siehe z.B. [51]).

Im folgenden Abschnitt werden wir für den Spezialfall der geometrischen Brownschen Bewegung explizite Formeln für die Verteilungsfunktion und die Erwartungswerte der Laufzeitextrema herleiten. Mit diesen Formeln könnten wir dann die Funktionalität der oben beschriebenen Routine zur numerischen Berechnung der Verteilungsfunktion überprüfen. Ferner kann man die Laufzeitextrema, ihre Verteilungsfunktionen und die entsprechenden Erwartungswerte simulieren. Wir haben bei einer umfangreichen Simulationsstudie (siehe [53]) festgestellt, dass sich die Laufzeitextrema der Returns eines Jump-Diffusions-Prozesses sich zwar von denen der geometrischen Brownschen Bewegung unterscheiden, die Charakteristiken aber nicht grundsätzlich verschieden sind.

7.3.1 Spezialfall der geometrischen Brownschen Bewegung

Wir werden die Erwartungswerte der Laufzeitextrema für den Spezialfall $\underline{p} \in D \setminus \hat{D}$ betrachten. Die Returns vereinfachen sich dann zu²

$$r_\tau = \ln \frac{S_\tau}{S_0} = \tilde{\mu}\tau + \sigma W_\tau = \left(\mu - \frac{\sigma^2}{2} \right) \tau + \sigma W_\tau.$$

Wir werden im Folgenden zur Vereinfachung das Running-Maximum mit Y_τ und das Running-Minimum mit Z_τ bezeichnen. Da die Pfade des stochastischen Prozesses $\mathbf{r} = (r_\tau, \tau \geq 0)$ stetig sind werden Supremum und Infimum der Laufzeitextrema tatsächlich angenommen, d.h. es gilt

$$\begin{aligned} Y_\tau &= \sup_{0 \leq t \leq \tau} r_t = \max_{0 \leq t \leq \tau} r_t \\ Z_\tau &= \inf_{0 \leq t \leq \tau} r_t = \min_{0 \leq t \leq \tau} r_t. \end{aligned}$$

Der folgende Satz aus [59, Corollary B.3.2.] liefert die Verteilungsfunktion des Running-Maximums Y_τ .

Satz 7.3.2 *Für jedes $c \geq 0$ ist die Verteilungsfunktion des Running-Maximums Y_τ durch*

$$F_{Y_\tau}(c) = \mathbb{P}(Y_\tau \leq c) = \Phi\left(\frac{c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) - e^{2\tilde{\mu}\sigma^{-2}c} \Phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) \quad (7.12)$$

gegeben.

²Für jeden Parametervektor $\underline{p} \in D \setminus \hat{D}$ gilt $\lambda(\mu_Y^2 + \sigma_Y^2) = 0$. Wegen (2.3) gilt dann $\lambda\nu = 0$, so dass aus (2.6) in diesem Spezialfall $\tilde{\mu} = \mu - \frac{\sigma^2}{2}$ folgt.

Folgerung 7.3.3 *Aus der Formel (7.12) erhalten wir durch Berechnung der Ableitung die Dichtefunktion von Y_τ mit*

$$f_{Y_\tau}(c) = \frac{1}{\sigma\sqrt{\tau}}\phi\left(\frac{c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) - \frac{2\tilde{\mu}}{\sigma^2}e^{2\tilde{\mu}c\sigma^{-2}}\Phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) + \frac{1}{\sigma\sqrt{\tau}}e^{2\tilde{\mu}c\sigma^{-2}}\phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right). \quad (7.13)$$

Theorem 7.3.4 *Der Erwartungswert des Running-Maximums ist durch*

$$\mathbb{E}Y_\tau = \begin{cases} \frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}}\Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}}\exp\left(-\frac{\tilde{\mu}^2\tau}{2\sigma^2}\right) - \frac{\sigma^2}{2\tilde{\mu}}, & \tilde{\mu} \neq 0 \\ 2\frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}}, & \tilde{\mu} = 0 \end{cases} \quad (7.14)$$

gegeben.

Beweis. Der Erwartungswert berechnet sich durch

$$\mathbb{E}Y_\tau = \int_0^\infty cf_{Y_\tau}(c)dc.$$

Wir betrachten zunächst Parametervektoren $\underline{p} \in D \setminus \hat{D}$ mit $\tilde{\mu} = 0$. Für diese vereinfacht sich die Dichtefunktion (7.13) zu

$$f_{Y_\tau}(c) = \frac{2}{\sigma\sqrt{\tau}}\phi\left(\frac{c}{\sigma\sqrt{\tau}}\right),$$

so dass für den Erwartungswert

$$\mathbb{E}Y_\tau = \frac{2}{\sigma\sqrt{\tau}} \int_0^\infty c\phi\left(\frac{c}{\sigma\sqrt{\tau}}\right)dc = 2\sigma\sqrt{\tau} \int_0^\infty z\phi(z)dz = 2\frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}}$$

folgt. Damit ist der zweite Teil obiger Aussage gezeigt. Sei nun im Weiteren $\tilde{\mu} \neq 0$. Mit (7.13) erhalten wir unmittelbar

$$\begin{aligned} \mathbb{E}Y_\tau &= \frac{1}{\sigma\sqrt{\tau}} \int_0^\infty c\phi\left(\frac{c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right)dc + \frac{1}{\sigma\sqrt{\tau}} \int_0^\infty ce^{2\tilde{\mu}c\sigma^{-2}}\phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right)dc \\ &\quad - \frac{2\tilde{\mu}}{\sigma^2} \int_0^\infty ce^{2\tilde{\mu}c\sigma^{-2}}\Phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right)dc. \end{aligned} \quad (7.15)$$

Den dritten Summanden von (7.15) berechnen wir mittels partieller Integration

$$\begin{aligned}
-\frac{2\tilde{\mu}}{\sigma^2} \int_0^\infty c e^{2\tilde{\mu}c\sigma^{-2}} \Phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) dc &= -\frac{2\tilde{\mu}}{\sigma^2} \left\{ \left[\frac{\sigma^2}{4\tilde{\mu}^2} e^{2\tilde{\mu}c\sigma^{-2}} (2\tilde{\mu}c - \sigma^2) \Phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) \right]_0^\infty \right. \\
&\quad \left. + \int_0^\infty \frac{\sigma}{4\tilde{\mu}^2\sqrt{\tau}} e^{2\tilde{\mu}c\sigma^{-2}} (2\tilde{\mu}c - \sigma^2) \phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) dc \right\} \\
&= -\frac{\sigma^2}{2\tilde{\mu}} \Phi\left(-\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) - \frac{1}{\sigma\sqrt{\tau}} \int_0^\infty c e^{2\tilde{\mu}c\sigma^{-2}} \phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) dc \\
&\quad + \frac{\sigma}{2\tilde{\mu}\sqrt{\tau}} \int_0^\infty e^{2\tilde{\mu}c\sigma^{-2}} \phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) dc,
\end{aligned}$$

so dass für den Erwartungswert

$$\begin{aligned}
\mathbb{E}Y_\tau &= \int_0^\infty \left\{ \frac{1}{\sigma\sqrt{\tau}} c \phi\left(\frac{c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) + \frac{\sigma}{2\tilde{\mu}\sqrt{\tau}} e^{2\tilde{\mu}c\sigma^{-2}} \phi\left(\frac{-c - \tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) \right\} dc - \frac{\sigma^2}{2\tilde{\mu}} \Phi\left(-\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) \\
&= -\frac{1}{4} \left[\frac{2\sqrt{2}\tilde{\mu}\sigma\tau e^{-\frac{(c-\tilde{\mu}\tau)^2}{2\sigma^2\tau}} - \sqrt{\pi\tau}(2\Phi\left(\frac{c-\tilde{\mu}\tau}{\sigma\sqrt{\tau}}\right) - 1)(2\tilde{\mu}^2\tau\sigma^2)}{\sqrt{\tau\pi}\tilde{\mu}} \right]_0^\infty - \frac{\sigma^2}{2\tilde{\mu}} \Phi\left(-\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) \\
&= \frac{2\tilde{\mu}^2\tau + \sigma^2}{4\tilde{\mu}} + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} e^{-\frac{\tilde{\mu}^2\tau}{2\sigma^2}} - \frac{2\tilde{\mu}^2\tau + \sigma^2}{4\tilde{\mu}} \left[2\Phi\left(-\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) - 1 \right] - \frac{\sigma^2}{2\tilde{\mu}} \Phi\left(-\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) \\
&= \frac{2\tilde{\mu}^2\tau + \sigma^2}{2\tilde{\mu}} \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} e^{-\frac{\tilde{\mu}^2\tau}{2\sigma^2}} - \frac{\sigma^2}{2\tilde{\mu}} \left[1 - \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) \right] \\
&= \frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}} \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} e^{-\frac{\tilde{\mu}^2\tau}{2\sigma^2}} - \frac{\sigma^2}{2\tilde{\mu}}
\end{aligned}$$

folgt und obige Aussage bewiesen ist \square

Lemma 7.3.5 *Der Erwartungswert $\mathbb{E}Y_\tau$ ist bezüglich $\tilde{\mu}$ stetig.*

Beweis. Durch Grenzübergang $\tilde{\mu} \rightarrow 0$ in (7.14) erhalten wir

$$\begin{aligned}
\lim_{\tilde{\mu} \rightarrow 0} \mathbb{E}Y_\tau &= \lim_{\tilde{\mu} \rightarrow 0} \left(\frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}} \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} e^{-\frac{\tilde{\mu}^2\tau}{2\sigma^2}} - \frac{\sigma^2}{2\tilde{\mu}} \right) \\
&= \lim_{\tilde{\mu} \rightarrow 0} \frac{\sigma^2 \left[\Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) - \frac{1}{2} \right]}{\tilde{\mu}} + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} \\
&= \lim_{\tilde{\mu} \rightarrow 0} \frac{\sigma^2 \left[\phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) \frac{\sqrt{\tau}}{\sigma} \right]}{1} + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} = 2 \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}},
\end{aligned}$$

so dass der Grenzwert $\lim_{\tilde{\mu} \rightarrow 0} \mathbb{E}Y_\tau$ und der Funktionswert von $\mathbb{E}Y_\tau$ für $\tilde{\mu} = 0$ übereinstimmen \square

Den Erwartungswert des Running-Minimums $\mathbb{E}Z_\tau$ können wir nun direkt von dem Erwartungswert des Running-Maximums ableiten.

Theorem 7.3.6 *Der Erwartungswert des Running-Minimums ist durch*

$$\begin{aligned}\mathbb{E}Z_\tau &= -\mathbb{E}Y_\tau + \tilde{\mu}\tau \\ &= \begin{cases} \tilde{\mu}\tau - \frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}} \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) - \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} \exp\left(-\frac{\tilde{\mu}^2\tau}{2\sigma^2}\right) + \frac{\sigma^2}{2\tilde{\mu}}, & \tilde{\mu} \neq 0 \\ -2\frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}}, & \tilde{\mu} = 0 \end{cases}\end{aligned}$$

gegeben.

Beweis. Wir überführen das Running-Minimum Z_τ mit

$$\begin{aligned}Z_\tau &= \min_{0 \leq t \leq \tau} r_t = \min_{0 \leq t \leq \tau} (\tilde{\mu}t + \sigma W_t) \\ &= -\max_{0 \leq t \leq \tau} (-\tilde{\mu}t - \sigma W_t) \stackrel{(d)}{=} -\max_{0 \leq t \leq \tau} (\tilde{\mu}t + \sigma \tilde{W}_t) \\ &= -\tilde{Y}_\tau\end{aligned}$$

in das -1 -Fache des Running-Maximums $\tilde{Y}_\tau := \max_{0 \leq t \leq \tau} \tilde{r}_t$ der Returns $\tilde{r}_t = \tilde{\mu}t + \sigma \tilde{W}_t$, da der Wiener-Prozess $\tilde{\mathbf{W}} = (\tilde{W}_t, t \in [0, \infty))$ mit $\tilde{W}_t = -W_t$ die gleiche Wahrscheinlichkeitsverteilung wie der Wiener-Prozess $\mathbf{W} = (W_t, t \in [0, \infty))$ besitzt. Für $\tilde{\mu} = -\mu \neq 0$ erhalten wir mit Theorem 7.3.4

$$\begin{aligned}\mathbb{E}Z_\tau &= -\mathbb{E}\tilde{Y}_\tau = -\left(\frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}} \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} e^{-\frac{\tilde{\mu}^2\tau}{2\sigma^2}} - \frac{\sigma^2}{2\tilde{\mu}}\right) \\ &= -\left\{-\frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}} \Phi\left(-\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} e^{-\frac{\tilde{\mu}^2\tau}{2\sigma^2}} + \frac{\sigma^2}{2\tilde{\mu}}\right\} \\ &= -\left\{\frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}} \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} e^{-\frac{\tilde{\mu}^2\tau}{2\sigma^2}} - \frac{\sigma^2}{2\tilde{\mu}}\right\} + \tilde{\mu}\tau = -\mathbb{E}Y_\tau + \tilde{\mu}\tau.\end{aligned}$$

Da für $\tilde{\mu} = -\mu = 0$ die Beziehung $\mathbb{E}\tilde{Y}_\tau = \mathbb{E}Y_\tau$ erfüllt ist, gilt dann ebenfalls $\mathbb{E}Z_\tau = -\mathbb{E}Y_\tau + \tilde{\mu}\tau$, so dass obige Aussage komplett bewiesen ist \square

Mit den Aussagen von Theorem 7.3.4 und Theorem 7.3.6 können wir nun die Maße (7.5) und (7.7) für den Spezialfall der geometrischen Brownschen Bewegung berechnen.

Folgerung 7.3.7 *Für das Vergleichsmaß ECAR des Maßes ACAR (vgl. (7.3)) gilt $ECAR(\tau) = \left(\mu - \frac{\sigma^2}{2}\right)\tau$. Das Vergleichsmaß ECBR des Maßes ACBR (vgl. (7.4)), das die Gegenbewegung einer Overreaction-Situation durch Mittlung aller bestmöglichen Returns nach den jeweiligen Indikationen qualifizierter 3-Tages Abwärtsbewegungen bewertet, berechnet sich für den Spezialfall der geometrischen Brownschen Bewegung wie folgt*

$$ECBR(\tau) = \begin{cases} \frac{\tilde{\mu}^2\tau + \sigma^2}{\tilde{\mu}} \Phi\left(\sqrt{\tau}\frac{\tilde{\mu}}{\sigma}\right) + \frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}} \exp\left(-\frac{\tilde{\mu}^2\tau}{2\sigma^2}\right) - \frac{\sigma^2}{2\tilde{\mu}}, & \tilde{\mu} \neq 0 \\ 2\frac{\sigma\sqrt{\tau}}{\sqrt{2\pi}}, & \tilde{\mu} = 0. \end{cases}$$

Ferner gilt

$$ECWR(\tau) = -ECBR(\tau) + \tilde{\mu} \tau$$

für das Vergleichsmaß von ACWR (vgl. (7.6)).

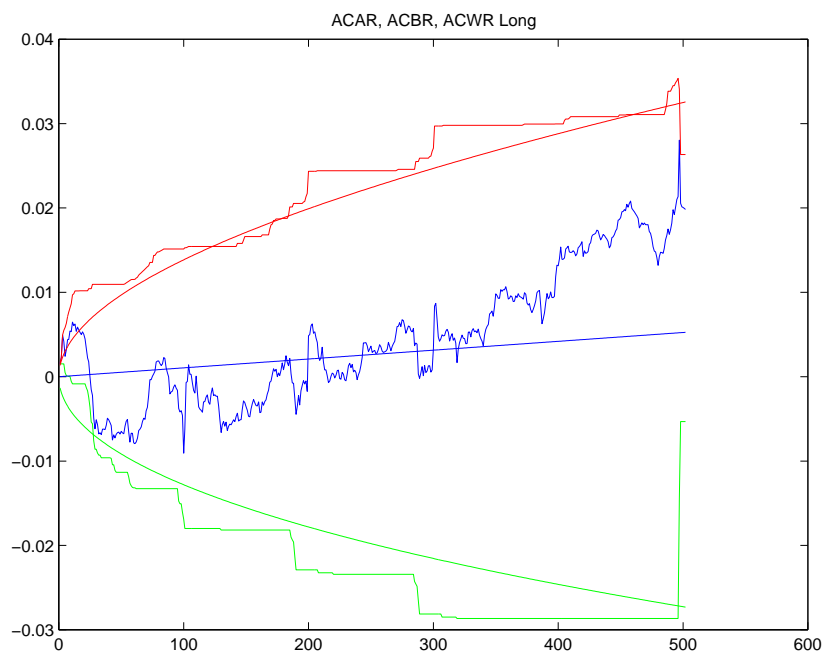


Abbildung 7.2: Overreaction-Maße beim DAX-Future 97/12 (5-Minuten-Takt)

Abbildung 7.2 zeigt alle Overreaction-Maße ACAR (blau), ACBR (rot) sowie ACWR (grün) beim DAX-Future mit Fälligkeit Dezember 1997 sowie die zugehörigen Erwartungswerte (glatte Linien) zum Vergleich. Dabei wurden nach jeder qualifizierten 3-Tages Abwärtsbewegung die Intraday-Futurekurse aller 5 Minuten für insgesamt 5 Handelstage ausgewertet. In der Grafik ist eine klare Abweichung des Maßes ACAR von der im Mittel zu erwartenden Entwicklung ECAR nach ca. 4 Handelstagen zu erkennen. Die Werte der bestmöglichen Returns ACBR liegen teilweise oberhalb der zu erwartenden Entwicklung, die sich im Mittel bei dem Kursmodell der geometrischen Brownschen Bewegung ergeben. Allerdings liegen in diesem Beispiel die Werte von ACWR auch unter den Werten von ECWR. Die Abweichungen zwischen den Kenngrößen ACBR und ECCR sowie ACWR und ECWR sind aber nicht so signifikant, wie die zwischen den Größen ACAR und ECAR.

Der Overreaction-Indikator sowie die Maße ACAR, ACBR und ACWR und deren Vergleichsfunktionen ECAR, ECCR und ECWR liefern eine umfangreiche Toolbox, um den Overreaction-Effekt näher zu untersuchen. Die von uns durchgeführten empirischen Studien mit den Kursen des DAX-Futures für den Spezialfall der geometrischen Brownschen Bewegung zeigen, dass der von uns vorgeschlagene Ansatz erfolgversprechend ist. Einerseits scheint der Overreaction-Indikator recht zuverlässig Overreaction-Situationen zu identifizieren, d.h. in den meisten Fällen sind den angezeigten Extrembewegungen

die vorhergesagten Gegenbewegungen gefolgt. Andererseits können die Gegenbewegungen mit den Maßen ACAR, ACBR und ACWR vernünftig analysiert und nachgewiesen werden. Der Vergleich dieser Maße mit den entsprechenden Erwartungswerten kann ferner Aufschluss über die Stärke der Gegenbewegung geben.

Abkürzungs- und Symbolverzeichnis

Symbol	Bedeutung
$\mathbf{S} = (S_t, t \in [0, \infty))$	stochastischer Prozess des Aktienkurses
S	Trajektorie des stochastischen Prozesses \mathbf{S}
$(\Omega, \mathfrak{F}, \mathbb{P})$	Wahrscheinlichkeitsraum
$\mathbf{W} = (W_t, t \in [0, \infty))$	Wiener-Prozess
$\mathbf{N} = (N_t, t \in [0, \infty))$	Poisson-Prozess mit Intensität λ
$\mathbf{N}^c = (N_t^c, t \in [0, \infty))$	zusammengesetzter Poisson-Prozess mit Intensität λ und Sprungamplituden $Y_j - 1$
$r_{x,x+\tau}$	Return vom Zeitpunkt x mit Lag τ
$\mathbf{r} = (r_{0,\tau}, \tau \in [0, \infty))$	Prozess der Returns
r^δ	empirische Returns
$\xi_1 \stackrel{(d)}{=} \xi_2$	Zufallsgrößen ξ_1 und ξ_2 besitzen die gleiche Verteilung
$X_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$	Konvergenz von $(X_n)_n$ gegen X in Wahrscheinlichkeit, d.h. $\forall \varepsilon > 0$ gilt $\lim_{n \rightarrow \infty} \mathbb{P}(X_n - X > \varepsilon) = 0$
\mathbb{I}	Indikatorfunktion
$\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)$	Parametervektor
\underline{D}	Definitionsbereich
$\mathbb{E}X$	Erwartungswert der Zufallsgröße X
$\ \cdot\ $	euklidische Norm
Φ	Verteilungsfunktion einer standardnormalverteilten Zufallsgröße
ϕ	Dichtefunktion einer standardnormalverteilten Zufallsgröße
φ	charakteristische Funktion der Returns
F	Verteilungsfunktion der Returns
f	Dichtefunktion der Returns
\tilde{f}	diskretisierte Dichtefunktion der Returns
$m_{\tau,k}$	k -tes Moment des Returns r_τ
$s_{\tau,k}$	k -te Semiinvariante des Returns r_τ
\mathbf{A}	Vorwärtsoperator
$\tilde{\mathbf{A}}$	diskretisierter Vorwärtsoperator
$\underline{\alpha}$	Regularisierungsparametervektor
α	skalärer Regularisierungsparameter

Literaturverzeichnis

- [1] C. M. Ahn, H. E. Thompson. Jump-diffusion processes and the term structure of interest rates. *The Journal of Finance*, XLIII(1):155–174, March 1988.
- [2] P. Albrecht, C. Kantar. Random walk oder mean reversion? *Kredit und Kapital*, 2:223–245, 2004.
- [3] T. G. Andersen, L. Benzoni, J. Lund. An empirical investigation of continuous-time equity return models. *The Journal of Finance*, LVII(3):1239–1284, June 2002.
- [4] L. Andersen, J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4:231–262, 2000.
- [5] D. D. Ang, R. Gorenflo, V. K. Le, D. D. Trong. *Moment Theory and Some Inverse Problems in Potential Theory and Heat Conduction*. Springer, Berlin, 2002.
- [6] S. Asmussen, J. Rosiński. Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Prob.*, 38:482–493, 2000.
- [7] J. Baumeister. *Stable Solution of Inverse Problems*. Vieweg, Braunschweig, 1987.
- [8] F. Black, M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 3(81):637–654, 1973.
- [9] M. Bremer, T. Hiraki, J. Sweeney. Predictable patterns after large stock price changes on the Tokyo stock exchange. *Journal of Financial and Quantitative Analysis*, 32(3):345–365, 1997.
- [10] T. F. Coleman, Y. Li. On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming*, 67(2):189–224, 1994.
- [11] T. F. Coleman, Y. Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.
- [12] R. Cont, P. Tankov. *Financial Modelling With Jump Processes*. Chapman & Hall/CRC, London, 2004.
- [13] R. Cont, P. Tankov. Non-parametric calibration of jump diffusion option pricing models. *The Journal of Computational Finance*, 7(3), 2004.

- [14] W. F. M. De Bondt, R. Thaler. Does the stock market overreact? *Journal of Finance*, 40(3):793–805, July 1985.
- [15] D. Duffie, J. P. K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, November 2000.
- [16] D. Duffie, K. J. Singleton. Simulated moments estimation of markov models of asset prices. *Econometrica*, 61(4):929–952, July 1993.
- [17] D. Düvelmeyer. Some stability results of parameter identification in a jump diffusion model. In J. vom Scheidt, editor, *Workshop-Tagungsband*, TU-Chemnitz, 2005. ISSN 1612-5665.
- [18] D. Düvelmeyer, B. Hofmann. Chances and limitations of a short-term overreaction effect - stochastic considerations and empirical results. In *PAMM - Proceedings in Applied Mathematics and Mechanics*, pages 459–460, 2003.
- [19] D. Düvelmeyer, B. Hofmann. Ill-posedness of parameter estimation in jump diffusion processes. In J. vom Scheidt, editor, *Workshop-Tagungsband*, pages 5–20, TU-Chemnitz, 2004. ISSN 1612-5665.
- [20] D. Düvelmeyer. Untersuchungen zu Chancen und Risiken von Anti-Trend-Strategien am Beispiel des DAX-Futures. Diplomarbeit, TU Chemnitz, Fakultät für Mathematik, http://www.tu-chemnitz.de/~dana/diplom_pdf_dd.zip, Januar 2001.
- [21] P. P. B. Eggermont. Maximum entropy regularization for Fredholm integral equations of the first kind. *SIAM J. Math. Anal.*, 24:1557–1576, 1993.
- [22] H. W. Engl, M. Hanke, A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, London, 2000.
- [23] H. W. Engl, K. Kunisch, A. Neubauer. Convergence rates for tikhonov regularization of non-linear ill-posed problems. *Inverse Problems*, 5:523–540, 1989.
- [24] B. Eraker. Do stock prices and volatility jump? *The Journal of Finance*, LIX(3):1367–1403, June 2004.
- [25] B. Eraker, M. Johannes, N. Polson. The impact of jumps in volatility and returns. *The Journal of Finance*, LVIII(3):1269–1300, June 2003.
- [26] S. N. Evans, P. B. Stark. Inverse problems as statistics. *Inverse Problems*, 18(4):R55–97, 2002.
- [27] E. F. Fama. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49:283–306, 1998.
- [28] G. M. Fichtenholz. *Differential- und Integralrechnung II*. Dt. Verl. d. Wissenschaften, Berlin, 1971.

- [29] A. K.-W. Fung, D. M. Mok, K. Lam. Intraday price reversals for index futures in the US and Hong Kong. *Journal of Banking & Finance*, 24:1179–1201, 2000.
- [30] A. R. Gallant, G. Tauchen. The relative efficiency of method of moments estimators. *Journal of Econometrics*, 92(1):149–172, September 1999.
- [31] C. Gaunt. Overreaction in the Australian equity market. *Pacific-Basin Finance Journal*, 8:375–398, 2000.
- [32] C. W. Groetsch. *Inverse Problems in the Mathematical Sciences*. Vieweg, Braunschweig, 1993.
- [33] J. Hadamard. *Lectures on the Cauchy Problem in Linear Partial Differential Equations*. Yale University Press, New Haven, 1923.
- [34] J. D. Hamilton. *Time Series Analysis*. Princeton University Press, New Jersey, 1994.
- [35] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6 No. 2, 2000.
- [36] B. Hofmann. *Regularization for applied inverse ill-posed problems*. Teubner, Leipzig, 1986.
- [37] B. Hofmann. An improved multi-parameter adjustment algorithm for inverse eigenvalue problems. *Zeitschrift für Analysis und ihre Anwendungen*, 7(6):511–518, 1988.
- [38] B. Hofmann. *Mathematik Inverser Probleme*. B. G. Teubner, Stuttgart, 1999.
- [39] B. Hofmann. Chances of strategies against the trend for DAX and DAX-Future based on a three-days indicator. Preprint TU-Chemnitz, Fakultät für Mathematik, 2000.
- [40] H. S. Holdgrün. *Analysis I*. Leins Verlag, Göttingen, 1998.
- [41] P. Honoré. Pitfalls in estimating jump diffusion models. Working Paper, University of Aarhus Graduate School of Business, www.hha.dk/~phs, 1998.
- [42] J. Hull, A. White. The pricing of options with stochastic volatilities. *Journal of Finance*, 42:281–300, June 1987.
- [43] J. C. Hull. *Options, Futures & Other Derivatives*. Prentice Hall International Editions, Upper Saddle River, NJ 07458, fourth edition, 2000.
- [44] G. Inglese. *Recent results in the study of the moment problems*. In: Theory and Practice of Geophysical Data Inversion (eds.: A. Vogel et al.). Vieweg Verlag, Braunschweig-Wiesbaden, 1992.

- [45] A. Irle. *Finanzmathematik, Die Bewertung von Derivaten*. Teubner, Stuttgart, 1998.
- [46] M. Johannes. The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance*, LIX(1):227–260, February 2004.
- [47] L. K. Jones. Approximation-theoretic derivation of logarithmic entropy principles for inverse problems and unique extension of the maximum entropy method to incorporate prior knowledge. *SIAM J. Appl. Math.*, 49:650–661, 1989.
- [48] P. Jorion. On jump processes in the foreign exchange and stock markets. *The Review of Financial Studies*, 1(4):427–445, 1988.
- [49] D. N. Ketcher, B. D. Jordan. Short-term price reversals following major price innovations: Additional evidence on market overreaction. *Journal of Economics and Business*, 46:307–323, 1994.
- [50] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems*. Springer, New York, 1996.
- [51] V. V. Kryzhniy. Direct regularization of the inversion of real-valued Laplace transforms. *Inverse Problems*, 19:573–583, 2003.
- [52] S. Kullback. *Information theory and statistics*. Dover Publications, Inc., Mineola, New York, 1997.
- [53] N. Lorenz. Simulationsstudien zu Laufzeitextrema bei Jump-Diffusions-Prozessen. Bakkalareusarbeit, TU-Chemnitz, Fakultät für Mathematik, 2004.
- [54] A. K. Louis. *Inverse und schlecht gestellte Probleme*. Teubner, Stuttgart, 1989.
- [55] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, March 1976.
- [56] R. C. Merton. *Continuous-Time Finance*. Blackwell Publishers Inc., Cambridge, 1992.
- [57] B. Meyer. *Der Overreaction-Effekt am deutschen Aktienmarkt*. Schriftenreihe der SGZ-Bank. Fritz Knapp Verlag, Frankfurt am Main, 1994.
- [58] R. Michaely, R. H. Thaler, K. L. Womack. Price reactions to dividend initiations and omissions: Overreaction or drift? *Journal of Finance*, 50(2):573–608, 1995.
- [59] M. Musiela, M. Rutkowski. *Martingale Methods in Financial Modelling*. Springer Verlag, second edition, 1997.
- [60] J. Naas. *Mathematisches Wörterbuch mit Einbeziehung der theoretischen Physik*. Teubner, Leipzig, 1984.
- [61] L. T. Nielson. *Pricing and Hedging of Derivative Securities*. Oxford University Press, New York, 1999.

- [62] W. R. Pestman. *Mathematical Statistics*. de Gruyter, Berlin, 1998.
- [63] P. Protter. *Stochastic Integration and Differential Equations*. Springer Verlag, 1992.
- [64] A. Rieder. *Keine Probleme mit Inversen Problemen*. Vieweg, Wiesbaden, 2003.
- [65] W. J. Runggaldier. Jump diffusion models. In: *Handbook of Heavy Tailed Distributions in Finance* (S. T. Rachev, ed.), Elsevier/North-Holland 2003, pp.169-209.
- [66] A. N. Shiryaev. *Probability*. Springer, New York, second edition, 1995.
- [67] A. N. Shiryaev. *Essentials of stochastic finance - Facts, Models, Theory*. World Scientific Publishing, 1999.
- [68] A. V. Skorohod. *Random Processes with Independent Increments*. Kluwer Academic Publishers, Dordrecht, 1991.
- [69] H.-J. Starkloff, D. Düvelmeyer, B. Hofmann. A note on uniqueness of parameter identification in a jump diffusion model. In J. vom Scheidt, editor, *Workshop-Tagungsband*, TU-Chemnitz, 2005. ISSN 1612-5665.
- [70] A. N. Tikhonov, V. Y. Arsenin. *Solutions of Ill-posed Problems*. V. H. Winston & Sons, Washington, D. C., 1977.
- [71] O. A. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.
- [72] G. Winkler. Analytische und numerische Untersuchungen des Modells von Heston zur Optionspreisberechnung. Diplomarbeit, TU Chemnitz, Fakultät für Mathematik, http://www-user.tu-chemnitz.de/~wgu/diplom/diplom_gw.pdf, März 2001.
- [73] R. Wunderlich. *Der Likelihoodquotiententest bei gruppierten Beobachtungen*. Dissertation, Technische Universität Chemnitz, Fakultät für Mathematik, 1992.
- [74] C. Zhou. A Jump Diffusion Approach to Modeling Credit Risk and Valuing Defaultable Securities. *Technical Report, Federal Reserve Board, Washington*, 1997.

Thesen

1. Die vorliegende Dissertation widmet sich dem inversen Problem der Parameterschätzung in Jump-Diffusions-Prozessen. Das hier betrachtete Kursmodell der Jump-Diffusions-Prozesse besteht dabei aus zwei Komponenten, einem Diffusionsteil und einem Sprungteil, die zusammen durch einen Parametervektor $\underline{p} = (\mu, \sigma, \lambda, \mu_Y, \sigma_Y)$ mit fünf Parametern spezifiziert werden. Die Parameter μ und σ bestimmen Drift und Volatilität des Diffusionsteils. Die Sprungintensität λ beeinflusst die Häufigkeit der Kurssprünge während die Parameter μ_Y und σ_Y die mittlere Höhe der Sprünge sowie deren Volatilität bestimmen. In vielen finanzmathematischen Anwendungen, z.B. der Optionspreisbestimmung, müssen die Parameter aus einer relativ kleinen Zahl von beobachteten Aktienkursen bestimmt werden.
2. In der Literatur werden verschiedene statistische Zugänge zur Parameterbestimmung diskutiert. Insbesondere gibt es Arbeiten, die auf Stabilitätsprobleme hinweisen, sofern die Parameter mit der herkömmlichen Maximum-Likelihood-Methode geschätzt werden. Tatsächlich existieren verschiedene Parameterkonstellationen, die zufallsbedingt ähnliche Trajektorien des Jump-Diffusions-Prozesses generieren können. So erzeugen Parametervektoren, die viele kleine Sprünge implizieren, ähnliche Trajektorien wie Parametervektoren, deren Parameter eine hohe Volatilität des Diffusionsteils verursachen. Daher werden in dieser Arbeit die statistischen Zugänge zur Parameterschätzung mit Techniken aus der Theorie inverser Probleme kombiniert.
3. Der in dieser Arbeit verfolgte Ansatz zur Parameterschätzung nutzt den Zusammenhang zwischen einer vom Parametervektor \underline{p} abhängigen Trajektorie des Jump-Diffusions-Prozesses und der Dichtefunktion $f_{r_\tau}(\cdot, \underline{p})$ der Returns r_τ . Der nichtlineare Vorwärtsoperator A , der jedem Parametervektor \underline{p} aus dem Definitionsbereich D die stetige Dichtefunktion $f_{r_\tau}(\cdot, \underline{p}) \in L^2(\mathbb{R})$ zuordnet, erweist sich als stetiger Operator. Allerdings ist der Operator A nicht für alle Parametervektoren $\underline{p} \in D = \{\underline{p} \in \mathbb{R}^5 : \sigma > 0, \lambda \geq 0, \sigma_Y \geq 0\}$ injektiv. Es zeigt sich, dass der Nichtinjektivitätsfall nur für Parametervektoren \underline{p} mit $\lambda(\sigma_Y^2 + \mu_Y^2) = 0$ eintritt, so dass der Operator A injektiv auf dem eingeschränkten Definitionsbereich $\hat{D} = \{\underline{p} \in D : \lambda(\sigma_Y^2 + \mu_Y^2) \neq 0\}$ ist. Daher ist jede Lösung $\underline{p} \in \hat{D}$ der Operatorgleichung $A(\underline{p}) = f$ eindeutig bestimmt. Ferner wird in dieser Arbeit gezeigt, dass für alle Lösungen $\underline{p} \in D \setminus \hat{D}$ die Diffusionsparameter μ und σ übereinstimmen.

4. Die Korrektheit der Operatorgleichung $A(\underline{p}) = f$, insbesondere deren Stabilität wurde ausführlich untersucht. Durch Beschränkung der Sprungparameter $\lambda \leq \lambda_{\max}$, $|\mu_Y| \leq \mu_{Y_{\max}}$ sowie $\sigma_Y \leq \sigma_{Y_{\max}}$ kann Stabilität im folgenden Sinne nachgewiesen werden: Falls eine Folge f_n im Raum $L^2(\mathbb{R})$ bzw. $C(\mathbb{R})$ gegen die Dichtefunktion f_0 konvergiert, so konvergieren die Diffusionsparameter der Urbilder $\underline{p}_n \in U(f_n)$ gegen die Diffusionsparameter der Parametervektoren $\underline{p}_0 \in U(f_0)$. Ferner konvergiert der Ausdruck $\lambda_n (\mu_{Y_n}^2 + \sigma_{Y_n}^2)$ mit den Sprungparametern λ_n , μ_{Y_n} und σ_{Y_n} der Urbilder $\underline{p}_n \in U(f_n)$ für $n \rightarrow \infty$ gegen den Ausdruck $\lambda_0 (\mu_{Y_0}^2 + \sigma_{Y_0}^2)$. Werden die Sprungparameter nicht beschränkt, so können Instabilitätsprobleme auftauchen, die durch eine asymptotische Nichtinjektivität verursacht werden.
5. Trotz der Stabilitätsaussagen treten bei der numerischen Lösung des diskretisierten Least-Squares Problems bereits bei kleinen Datenfehlern $\|f - f^\delta\| \leq \delta$ Inkorrektkeitsphänomene auf. Zwischen der exakten Lösung und der geschätzten Lösung \underline{p}^δ sind teilweise erhebliche Abweichungen festzustellen. Diese Inkorrektkeitsphänomene werden in einer umfangreichen numerischen Fallstudie nachgewiesen und illustriert.
6. Es wird ein Multiparameter-Regularisierungszugang vorgestellt, um diese Instabilitätsprobleme zu überwinden. Bei der vorgeschlagenen Regularisierung wird das Minimum des ursprünglichen Least-Squares Problems nicht mehr für alle Parametervektoren $\underline{p} \in D$ bestimmt, sondern nur für Parametervektoren deren Semiinvarianten sich in der Umgebung der empirischen Semiinvarianten befinden. Der Regularisierungszugang führt auf ein nichtlineares und nichtkonvexes Optimierungsproblem, dessen Optimallösungen geeignete Approximationen des exakten Parametervektors sind. Das Lagrangefunktional dieser Optimierungsaufgabe entspricht im Wesentlichen einem Regularisierungsfunktional bei der Tichonov-Regularisierung. Indem für verschiedene Regularisierungsparametervektoren jeweils die Lösung des verwandten regularisierten Problems berechnet wird, kann durch eine Fixpunktiteration die Optimallösung angenähert werden. Es wird ein entsprechender Iterationsalgorithmus zur Bestimmung der Optimallösung vorgestellt. In der Arbeit wird nachgewiesen, dass jeder Fixpunkt, gegen den dieser Algorithmus konvergiert, einer Optimallösung entspricht. Zahlreiche numerische Rechnungen mit simulierten Daten zeigen, dass der Algorithmus in praktischen Situationen konvergiert und die erhaltenen Lösungen sehr gute Approximationen der exakten Lösung darstellen. Das Problem der Wahl numerischer Parameter wird anhand einer Fallstudie diskutiert.
7. Als ein konkretes Anwendungsbeispiel zur Parameterschätzung wird die Indikation von Overreaction-Effekten vorgestellt, welche auf der Berechnung von Wahrscheinlichkeiten extremer Kursbewegungen beruht. Der Overreaction-Effekt wird durch irrationales Handeln vieler Marktteilnehmer in ungewöhnlichen und unerwarteten Situationen verursacht. Extremen Kursbewegungen folgen dann entsprechende Gegenbewegungen. In dieser Arbeit wurden Maße für den Overreaction-Effekt eingeführt. Es zeigt sich, dass durch Vergleich dieser Maße mit den Erwartungswerten der Laufzeitextrema der Returns Aufschluss über das Ausmaß

der Gegenbewegung erhalten werden kann. Es wird ein allgemeiner Zugang zur Berechnung dieser Erwartungswerte für das Kursmodell der Jump-Diffusions-Prozesse vorgestellt. Für den Spezialfall der geometrischen Brownschen Bewegung werden explizite Formeln hergeleitet.

Lebenslauf

Persönliche Daten

Name:	Dana Düvelmeyer
Geburtsdatum:	20. August 1977
Geburtsort:	Greiz
Familienstand:	ledig

Schul Ausbildung

Sept.1984-Aug.1985	Goethe-Oberschule in Greiz
Sept.1985-Aug.1990	Käthe-Kollwitz-Oberschule in Karl-Marx-Stadt
Sept.1990-Aug.1991	Obere Luisenschule in Chemnitz
Sept.1991-Jul.1992	Spezialschule Eissport in Chemnitz
Aug.1993-Jun.1996	Sportgymnasium in Chemnitz
	Abschluss: Abitur (Durchschnitt: 1.5)

Studium

Okt.1996-Januar 2001	Studium der Wirtschaftsmathematik an der TU-Chemnitz, Abschluss als Diplom-Wirtschaftsmathematikerin (Prädikat: „mit Auszeichnung“)
----------------------	---

Wissenschaftlicher Werdegang

seit März 2001	wissenschaftliche Mitarbeiterin an der Fakultät für Mathematik der TU-Chemnitz
----------------	---

Erklärung

Ich erkläre an Eides Statt, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Chemnitz, den 20. Januar 2005